

XX PA (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.
XX PI Hiscott J, Lin R;
XX DR WPI; 1999-620201/53.
XX DR N-PSDB; AAX30993.
XX Carboxy-terminus modified highly active forms of interferon regulatory
PT factor proteins used for the treatment of viral infections
XX Claim 14; Fig 12; 93pp; English.
XX The present protein sequence is the interferon regulatory factor-7
CC (IRF-7) 2D protein, that is modified in the transactivation domain. The
CC serine and threonine residues in the carboxy terminus are modified
CC post-translationally, by phosphorylation, following sendai virus
CC infection. The modified IRF-7, substituted with aspartic acid, functions
CC as a activator of promoters containing ISRE regulatory elements and
CC stimulation of DNA binding and transcriptional activity. IRF-7 protein
CC shows highest homology to IRF-3. The modified IRF proteins are used for
CC the treatment of viral infections like, influenza, herpes or HIV
CC infection. They may also be used to activate a cytokine gene, in cancer
CC treatment or to modify a target cell of an organism.
XX SQ Sequence 503 AA;
Query Match 100.0%; Score 2731; DB 20; Length 503;
Best Local Similarity 100.0%; Pred. No. 1.6e-197;
Matches 503; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALAPERAAPRVLFGEWLLGEISSGCGEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAPRVLFGEWLLGEISSGCGEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWPPSSRGSGGPPPEAETAEARAGWKTNFRFCALRSTRFRVWLRDNGSDPADPHK 120
DB 61 KAWAVARGWPPSSRGSGGPPPEAETAEARAGWKTNFRFCALRSTRFRVWLRDNGSDPADPHK 120
QY 121 VVALSRELCHWREGPGTDTQTEAEAPAAVPPQGGPPGFLAHTAGLQAGPLPAPAGDKG 180
DB 121 VVALSRELCHWREGPGTDTQTEAEAPAAVPPQGGPPGFLAHTAGLQAGPLPAPAGDKG 180
QY 181 DLLLQAVQSCSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
DB 181 DLLLQAVQSCSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
QY 241 ETTTSPGPQPAALTGEEAAAPESPHQAEPLYLSPSPSACTAVQSPSCALDVTIMYKGRV 300
DB 241 ETTTSPGPQPAALTGEEAAAPESPHQAEPLYLSPSPSACTAVQSPSCALDVTIMYKGRV 300
QY 301 LQKVVGHPSCFTLYGPPDPVAVRATDQQAFAFSPAPLDPQKQRYTEELLRHVAPGLHLE 360
DB 301 LQKVVGHPSCFTLYGPPDPVAVRATDQQAFAFSPAPLDPQKQRYTEELLRHVAPGLHLE 360
QY 361 LRGPQLWARMGKCKVYWEVGGPPGSGASPSGTACLLPRNCDTPIFDFRVFFQSLVEFRAR 420
DB 361 LRGPQLWARMGKCKVYWEVGGPPGSGASPSGTACLLPRNCDTPIFDFRVFFQSLVEFRAR 420
QY 421 QRRGSPRYTILYFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSDLDL 480
DB 421 QRRGSPRYTILYFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSDLDL 480
QY 481 CLSSANSLYDDIECFMELQPA 503
DB 481 CLSSANSLYDDIECFMELQPA 503
RESULT 2
AAE09329
ID AAE09329 standard; Protein; 503 AA.
XX
AC AAE09329;

XX 19-NOV-2001 (first entry)
XX Human intracellular regulatory molecule, KWC02.
DE Human intracellular regulatory molecule, KWC02.
XX Human; intracellular regulator; cell division; proliferation; therapy;
KW cancer; infection; wound; developmental abnormality; metabolic problem;
KW cytostatic; antibacterial; vulnery; transcription factor; KWC02;
KW interferon response factor; IRF.
XX Homo sapiens.
OS US6274312-B1.
XX 14-AUG-2001.
PD 10-DEC-1997; 97US-0999774.
XX 11-DEC-1996; 96US-0032818.
PR (SCHE) SCHERING CORP.
PA Gish KC, Seghezzi W, Shanahan F, Lees EM, McClanahan TK;
XX WPI; 2001-535086/59.
PI N-PSDB; AAD16105.
XX New genes encoding intracellular regulatory molecules, useful for
PT regulating cell division and proliferation (e.g. tumor cells),
PT particularly for treating cancer, infections, wounds, or developmental
PT or metabolic abnormalities
XX Example 2; Column 33-38; 62pp; English.
PS The present sequence is a human intracellular regulatory
XX molecule, KWC02. The KWC02, an interferon response factor (IRF)
CC homologue, is a transcription factor. The polynucleotides encoding
CC intracellular regulatory molecules are useful for regulating cell
CC division and proliferation of various cell types, including tumour
CC cells. Specifically, they are also useful for treating cancer,
CC infections, wounds, developmental abnormalities or metabolic problems.
XX Sequence 503 AA;
Query Match 98.7%; Score 2695; DB 22; Length 503;
Best Local Similarity 99.0%; Pred. No. 8.4e-195;
Matches 498; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 MALAPERAAPRVLFGEWLLGEISSGCGEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAPRVLFGEWLLGEISSGCGEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWPPSSRGSGGPPPEAETAEARAGWKTNFRFCALRSTRFRVWLRDNGSDPADPHK 120
DB 61 KAWAVARGWPPSSRGSGGPPPEAETAEARAGWKTNFRFCALRSTRFRVWLRDNGSDPADPHK 120
QY 121 VVALSRELCHWREGPGTDTQTEAEAPAAVPPQGGPPGFLAHTAGLQAGPLPAPAGDKG 180
DB 121 VVALSRELCHWREGPGTDTQTEAEAPAAVPPQGGPPGFLAHTAGLQAGPLPAPAGDKG 180
QY 181 DLLLQAVQSCSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
DB 181 DLLLQAVQSCSLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
QY 241 ETTTSPGPQPAALTGEEAAAPESPHQAEPLYLSPSPSACTAVQSPSCALDVTIMYKGRV 300
DB 241 ETTTSPGPQPAALTGEEAAAPESPHQAEPLYLSPSPSACTAVQSPSCALDVTIMYKGRV 300
QY 301 LQKVVGHPSCFTLYGPPDPVAVRATDQQAFAFSPAPLDPQKQRYTEELLRHVAPGLHLE 360
DB 301 LQKVVGHPSCFTLYGPPDPVAVRATDQQAFAFSPAPLDPQKQRYTEELLRHVAPGLHLE 360
QY 361 LRGPQLWARMGKCKVYWEVGGPPGSGASPSGTACLLPRNCDTPIFDFRVFFQSLVEFRAR 420

Db 361 LRGPOLWARMGCKVYEVGGPGSASPSTPACLLPRNCDTPIDFRVFFOELVEFRAR 420
QY 421 QRRGSPRYTYILGFGQDLISAGRPKEKSLVLVKLEPWLCKRVHLEGTQREGVSSLDSSLDL 480
Db 421 QRRGSPRYTYILGFGQDLISAGRPKEKSLVLVKLEPWLCKRVHLEGTQREGVSSLDSSLSL 480
QY 481 CLSSANSYDDIECFIMELEQPA 503
Db 481 CLSSANSYDDIECFIMELEQPA 503

RESULT 3
AA15104
ID AA15104 standard; Protein; 542 AA.
XX
AC AA15104;
DT 25-JAN-2000 (first entry)
XX
DE Chimeric protein IRP-7 (1-246)/IRP-3 (132-427).
KW Interferon regulatory factor; IRP; chimeric protein; serine; threonine;
KW carboxy terminus; amino terminal domain; aspartic acid; phosphorylation;
KW post-translational modification; sendai virus; cancer treatment; herpes;
KW PRDI/PRDIII; promoter; ISRE regulatory element; stimulation; activate;
KW activator; DNA binding; transcriptional activity; viral infection;
KW proteasome mediated degradation; influenza; HIV infection; cytokine gene;
KW target cell.
XX
OS Homo sapiens.
OS Synthetic.

Key Location/Qualifiers
FT Region 1..246
FT /note= "Corresponds to modified amino-terminal domain
FT of IRP-7 protein"
FT Region 247..542
FT /note= "Corresponds to modified carboxy-terminus of
FT IRP-3 protein"
FT Misc-difference 511
FT /note= "Wild type Ser replaced with Asp"
FT Misc-difference 513
FT /note= "Wild type Ser replaced with Asp"
FT Misc-difference 517
FT /note= "Wild type Ser replaced with Asp"
FT Misc-difference 519
FT /note= "Wild type Ser replaced with Asp"
FT Misc-difference 520
FT /note= "Wild type Ser replaced with Asp"
XX WO9951737-A1.
XX
PD 14-OCT-1999.
XX
XX
PF 07-APR-1999; 99WO-CA00314.
XX
PR 07-APR-1998; 98CA-2234588.
XX
PA (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.
XX
PI Hiscott J, Lin R;
XX
PI WPI; 1999-620201/53.
DR N-PSDB; AAX90994.
XX
DR Carboxy-terminus modified highly active forms of interferon regulatory
PT factor proteins used for the treatment of viral infections
XX
XX Claim 11; Fig 13; 93pp; English.
PS
XX The present sequence is the chimeric protein comprising, residues 1-246
CC from the modified amino-terminal domain of interferon regulatory factor,

CC IRP-7 and 132-427 residues from the carboxy-terminus of modified IRP-3
CC (SD) protein. The serine and threonine residues are post-translationally
CC modified by phosphorylation, following sendai virus infection. The
CC modified IRP, substituted with aspartic acid, functions as a strong
CC activator of promoters containing ISRE and PRDI/PRDIII regulatory
CC elements, stimulation of DNA binding and transcriptional activity and
CC proteasome mediated degradation. The modified IRP sequences are used for
CC the treatment of viral infections like, influenza, herpes or HIV
CC infection. They may also be used to activate a cytokine gene, in cancer
CC treatment or to modify a target cell of an organism.
XX
SQ Sequence 542 AA;

Query Match 57.4%; Score 1567.5; DB 20; Length 542;
Best Local Similarity 59.7%; Pred. No. 7.7e-110;
Matches 327; Conservative 38; Mismatches 116; Indels 67; Gaps 12;
QY 1 MALAPERAAPRVLPGEWLLGIISSCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60
Db 1 MALAPERAAPRVLPGEWLLGIISSCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWPPSSRGSGGPPPEAETAERAGWKTFRCALRSTRFRVMLRDNSSGDPADPHK 120
Db 61 KAWAVARGWPPSSRGSGGPPPEAETAERAGWKTFRCALRSTRFRVMLRDNSSGDPADPHK 120
QY 121 VYALSRELWCWREGPGTQDTEAEAPAAVPPQGGPPGPFPLATHAGLQAPGPPAPAGDKG 180
Db 121 VYALSRELWCWREGPGTQDTEAEAPAAVPPQGGPPGPFPLATHAGLQAPGPPAPAGDKG 180
QY 181 DILLQAVQOOSCLADHLLTASWGADVPVTKAPGEGOGELPLTGACAGGGLPAGELYGMVAV 240
Db 181 DILLQAVQOOSCLADHLLTASWGADVPVTKAPGEGOGELPLTGACAGGGLPAGELYGMVAV 240
QY 241 ETTTSS-----PGPOPAALTTCGAAAPESP HQAEPYLSPS----- 274
Db 241 ETTTSPSTQEDILDELLGNWVLAFLPDGPPSL-----AVAPECPQ--PLRSFLDNP 294
QY 275 -----PSACTAVQEPSG---ALDVTIMYKGRVTLQKVGHPSCTFLYGPDPAPVAT 324
Db 295 TFPNPLGSENPLKRLVPGSEWEPEVTAFYRGVQVFOQTISCPEGLRVG-SEVGDRTL 353
QY 325 DPQVAFPSPA-ELPDQKQLAYTELLRHVAPGLHLELPGPOLWARRMCKKCVYEVG-- 381
Db 354 PGWPTLPDPGMSLTDGVMYSYRVHVLSCGLGGLALWRAGQWLVAQRLGHCHTYWAVSEE 413
QY 382 -----GPGSASPSPTACLLPRNCDTPIDFRVFFOELVEFRARQRGSPRYTYILGF 434
Db 414 LLPNSGCHGPDSE-----VPKDKGCVFDLGFIVDLITFTTEGSGR-SRYALMFCV 463
QY 435 GQDLISAGRPKEKSLVLVKLEPWLCKRVHLEGTQREGVSSLDSSLDLCLSSANS- 491
Db 464 GESWPQDQFWTKRLVMVKVPTCLRALVEMARVGGASSLENT-VDLHIDNDHPDLDDDDQ 522
QY 492 IECFLMEL 499
Db 523 YKAVLQDL 530

RESULT 4
AAW78986
ID AAW78986 standard; Protein; 427 AA.
XX
AC AAW78986;
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1648.
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX

Db 64 YVP-----GRDKPDLT-----WKNFRSALNRKGLRLAEDRSKDPHPKHYEFV----- 110
Qy 130 WREGGTDQTEAEAPAAVPPQGGPPGFLAHTHAGLQAPGLPAPAGDKGDLQLLAQVQQ 189
Db 111 -NSGVG---DFSQPDTSPTDNGG-----GSTD-----TQ 136
Qy 190 SCLADHLTASGADPVTAPKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSPGQ 249
Db 137 EDILDELL-GNMVLAPL-----DFGPP-----SLAVAPEPCQ 169
Qy 250 PAALTGGAAPESPHQAEPLYSPSACTAVQEP-SPGALDVTIMYKGRVTLQKVGH 308
Db 170 PLSPSLNDPT-----FPNLGSENPLKRLLVPEGEWEFEVTFYGRQVFOQTISCP 223
Qy 309 SCTFLYGPDPFAVRATDQQVAFPPSPA-ELPDQKQLRYTEELLRHVAPGLHLELRGPQLW 367
Db 224 EGLRLVG-SEVGDRTLPGWVTLDPGMSLTDGVMYSVRHVLSCGLGGLALWRAGQLW 282
Qy 368 ARMGKCKVYEVG-----GPGSASPTPACILPRNCDTPIPFPRVFFQELVEFR 418
Db 283 AQRLGHCHTYWAVSEBELLPSNGHGDGE-----VPKKEGGVFDLGPPIVDLITFT 333
Qy 419 ARQREGSPRYTILYFGODLSAGRPKEKSLVLKLEPWLRCVHLRGTOREGVSSLDSD 478
Db 334 EGSGR-SPRYALWFCVGSWPQDQPTWKRLVMKVVPTCLRALVEMARVGGASLENT-V 391
Qy 479 DCLSSANSLS---YDDIECFMEL 499
Db 392 DLHIDNHPLDLDDQYKAYLQDL 415
RESULT 6
ID AAY15102
XX AAY15102 standard; Protein; 427 AA.
AC AAY15102;
XX
DT 25-JAN-2000 (first entry)
XX
DE Modified Interferon Regulatory Factor-3 protein (IRF-3).
XX
KW Interferon regulatory factor; IRF-3; transactivation domain; serine;
KW threonine; carboxy terminus; aspartic acid; mutant; phosphorylation;
KW post-translational modification; sendai virus; cancer treatment; herpes;
KW activator; promoter; PRDI/PRDII; ISRE regulatory element; stimulation;
KW DNA binding; transcriptional activity; transcriptional co-activator;
KW CBP/p300; proteasome mediated degradation; viral infection; influenza;
KW HIV infection; activate; cytokine gene; target cell.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 396 /note= "wild type Ser replaced with Asp"
FT FT
FT Misc-difference 398 /note= "wild type Ser replaced with Asp"
FT FT
FT Misc-difference 402 /note= "wild type Ser replaced with Asp"
FT FT
FT Misc-difference 404 /note= "wild type Ser replaced with Asp"
FT FT
FT Misc-difference 405 /note= "wild type Ser replaced with Asp"
FT FT
XX WO951737-A1.
XX
XX 14-OCT-1999.
XX
XX 07-APR-1999; 99WO-CA00314.
XX
XX 07-APR-1998; 98CA-2234588.
XX
XX (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.

XX Hiscott J, Lin R;
XX WPI; 1999-620201/53.
DR N-PSDB; AAX90992.
XX
FT Carboxy-terminus modified highly active forms of interferon regulatory
FT factor proteins used for the treatment of viral infections -
XX
PS Claim 8; Fig 10; 93pp; English.
XX
CC The present protein sequence is the interferon regulatory factor-3
CC (IRF-3)5D protein, that is modified in the transactivation domain. The
CC serine and threonine residues in the carboxy terminus are modified
CC post-translationally, by phosphorylation, following sendai virus
CC infection. The modified IRF-3, substituted with aspartic acid, functions
CC as a strong activator of promoters containing PRDI/PRDII or ISRE
CC regulatory elements, stimulation of DNA binding and transcriptional
CC activity, association of IRF-3 with the transcriptional co-activator
CC CBP/p300 and proteasome mediated degradation. The modified IRF proteins
CC are used for the treatment of viral infections like, influenza, herpes
CC or HIV infection. They may also be used to activate a cytokine gene, in
CC cancer treatment or to modify a target cell of an organism.
XX
SQ Sequence 427 AA;
Query Match 16.5%; Score 450; DB 20; Length 427;
Best Local Similarity 27.8%; Pred. No. 8.6e-26;
Matches 140; Conservative 63; Mismatches 193; Indels 108; Gaps 19;
Qy 10 PRVLFGWLLGEISSGCGYEGLOWLDEARTCRVRPKHFKDLSADARIPKAWAVARG 69
Db 6 PRIL--PWLVSQDLQGLQEGVAMVWNSKTRTRIPMKHGLRQDAQQEDFGIFQAWAETGA 63
Qy 70 WPPSSRGGPPPEAETAERAGWKTFRCAALSTRFVMLRNSGDPADPHKVVYALSREL 129
Db 64 YVP-----GRDKPDLT-----WKNFRSALNRKGLRLAEDRSKDPHPKHYEFV----- 110
Qy 130 WREGGTDQTEAEAPAAVPPQGGPPGFLAHTHAGLQAPGLPAPAGDKGDLQLLAQVQQ 189
Db 111 -NSGVG---DFSQPDTSPTDNGG-----GSTD-----TQ 136
Qy 190 SCLADHLTASGADPVTAPKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSPGQ 249
Db 137 EDILDELL-GNMVLAPL-----DFGPP-----SLAVAPEPCQ 169
Qy 250 PAALTGGAAPESPHQAEPLYSPSACTAVQEP-SPGALDVTIMYKGRVTLQKVGH 308
Db 170 PLSPSLNDPT-----FPNLGSENPLKRLLVPEGEWEFEVTFYGRQVFOQTISCP 223
Qy 309 SCTFLYGPDPFAVRATDQQVAFPPSPA-ELPDQKQLRYTEELLRHVAPGLHLELRGPQLW 367
Db 224 EGLRLVG-SEVGDRTLPGWVTLDPGMSLTDGVMYSVRHVLSCGLGGLALWRAGQLW 282
Qy 368 ARMGKCKVYEVG-----GPGSASPTPACILPRNCDTPIPFPRVFFQELVEFR 418
Db 283 AQRLGHCHTYWAVSEBELLPSNGHGDGE-----VPKKEGGVFDLGPPIVDLITFT 333
Qy 419 ARQREGSPRYTILYFGODLSAGRPKEKSLVLKLEPWLRCVHLRGTOREGVSSLDSD 478
Db 334 EGSGR-SPRYALWFCVGSWPQDQPTWKRLVMKVVPTCLRALVEMARVGGASLENT-V 391
Qy 479 DCLSSANSLS---YDDIECFMEL 499
Db 392 DLHIDNHPLDLDDQYKAYLQDL 415
RESULT 7
ID ABP42913
XX ABP42913 standard; Protein; 175 AA.
AC ABP42913;
XX

Db 26 QWLIDQIDSGKYPGLWNEEKSVERIPWKHAGKQDYNREEDAAALFKAWALFKGXF----- 81
QY 75 RGGGPPPEAETAERAGWKTNFRFCALSTRFRFVMLRDNSG-DPADPHKYVALSRELWCWREG 133
Db 82 REGIDKPPDPT-----WKTRLCALNKSNDFFELVERSOLDISDPYKVVRI----- 127
QY 134 PGTDTQTEAEAPAAVPPPGGPPGP---FLAHTHAGLQAPGLPAPAGDKGDLILLQAVQOS 190
Db 128 -----VPEGAKKGAKQLTDDTQMAMGHYPMTAP---YGSLLPAQOQVHNY 169
QY 191 CLADHLLTASWAGDVPVTKAPGEGOGELPLTGACAG-----GGLP-----AGELYGWAWE 241
Db 170 WMPPH--DRSW-RDYAPDOSHPEIYQCFVTGPRGHHWQGPSCEGCOVGTTFYACAPP 226
QY 242 TTPSPG-PQPAALTTGEAAAPESPHQAEPYLSFSPSACTAVQEPSPGALDVTIMYKGRV 300
Db 227 ESQAFGPIEPIESIRSAEALA-----LSDC-----RLHICLYYREILV 263
QY 301 LQKVVGHP-SCTFLYGGPPDPAVRATDPOQVAFPPSPAEPLDQKQLRYTELLRHVAPGLH 359
Db 264 KELTTTSPGCRISHG---HTYDVSNLDQVLPFPY---DDNGQRKNIEKLLSHLERGLV 317
QY 360 ELRGQQLWARRMGKCKVYVEGVPGPSAS--PSTPACLLPRNCPTPIDFRVFFQELVEF 417
Db 318 WMAPDGLYAKRLCOSRIYWD--GPLALCSDRNK-----LERDQTKCLFTQOFLSELOV 371
QY 418 RARQRGSPRYTIYLGFGQDLISAGRPKEKSLVLKLEPWLCL 459
Db 372 -AHGGRAPRFQVTLFCGEEFPDPO-RQRKLITAHVEPLLAR 411

RESULT 9

AAW99427
ID AAW99427 standard; Protein; 450 AA.
XX
AC AAW99427;
XX
DT 20-JAN-1997 (first entry)
XX
DE Human lymphocyte specific interferon regulatory factor.
XX
KW Lymphocyte specific interferon regulatory factor; LSIRF; IRF-3.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 163..164
FT /note="the Double Q form of hLSIRF contains an
FT additional glutamine residue between
FT residues 163 and 164 of the Single Q form"
XX
PN WO9632477-A1.
XX
PD 17-OCT-1996.
XX
PF 12-APR-1996; 96WO-CA00231.
XX
PR 03-APR-1996; 96US-0611280.
PR 14-APR-1995; 95US-0422733.
XX
PA (AMGE-) AMGEN CANADA INC.
XX
PI Grossman A, Matsuyama T, Richardson CD;
XX
XX WPI; 1996-477128/47.
DR N-PSDB; AAT41706.
XX
XX New genes for murine lymphocyte specific interferon regulatory
PT factor - used for modulation of lymphocyte activation and
PT proliferation
XX
PS Claim 1; Page 65-66; 92pp; English.
XX

CC Novel Single Q or Double Q forms of human lymphocyte-specific
CC interferon regulatory factor (LSIRF or ISF-3) (AAW99426) bind to
CC interferon-stimulated response elements in DNA. The amino acid
CC sequence of the Single Q form was deduced from a cDNA clone (AAT41706)
CC isolated from a human lymphocyte cDNA library. Recombinant LSIRF
CC can be produced in transfected prokaryotic or eukaryotic host
CC cells. LSIRF polypeptides are useful as targets for therapeutic
CC cpds. used to regulate lymphocyte activation. By blocking LSIRF
CC activity it is possible to suppress lymphocyte activation in
CC response to certain environmental stimuli.
XX
SQ Sequence 450 AA;
Query Match 13.2%; Score 361.5; DB 17; Length 450;
Best Local Similarity 27.2%; Pred. No. 4.3e-19;
Matches 135; Conservative 70; Mismatches 194; Indels 97; Gaps 22;
QY 16 EWLLEISGGYEGQLWLEDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGWPPSP 74
Db 26 QWLIDQIDSGKYPGLWNEEKSVERIPWKHAGKQDYNREEDAAALFKAWALFKGXF----- 81
QY 75 RGGGPPPEAETAERAGWKTNFRFCALSTRFRFVMLRDNSG-DPADPHKYVALSRELWCWREG 133
Db 82 REGIDKPPDPT-----WKTRLCALNKSNDFFELVERSOLDISDPYKVVRIVPEGAKK-- 134
QY 134 PGTDTQTEAEAP--AAVPPPGGPPGPFL-AHTHAGLQAPGLPAPAGDKGDLILLQAVQOS 190
Db 135 -GAKQLTLEDPQMSMSHPYTWTTTPYPSLPAQVHNYMMP----- 172
QY 191 CLADHLLTASWAGDVPVTKAPGEGOGELPLT-----GACAGGCLPAGELYGMVAV 240
Db 173 -----LDRSW-RDYVDPDPHPPEIPYQCFMTGPRGHHWQGPSCEGCOV-TGTFYACAP 224
QY 241 ETTPSPG-PQPAALTTGEAAAPESPHQAEPYLSFSPSACTAVQEPSPGALDVTIMYKGR 299
Db 225 PESQAFGVTPESIRSAEALA-----FSDC-----RLHICLYYREIL 261
QY 300 VLQKVVGHP-SCTFLYGGPPDPAVRATDPOQVAFPPSPAEPLDQKQLRYTELLRHVAPGLH 358
Db 262 VKELTTSSPEGCRISHG---HTYDASNLDQVLPFPY---EDNGQRKNIEKLLSHLERGV 315
QY 359 LELRGQQLWARRMGKCKVYVEGVPGPSAS--PSTPACLLPRNCPTPIDFRVFFQELVE 416
Db 316 LMAPDGLYAKRLCOSRIYWD--GPLALCSDRNK-----LERDQTKCLFTQOFLSELOA 369
QY 417 FRARQRGSPRYTIYLGFGQDLISAGRPKEKSLVLKLEPWLCLRVHLEGTRGVSLSLSS 476
Db 370 P-AHGRSLPRFQVTLFCGEEFPDPO-RQRKLITAHVEPLLARQLYFFAQNSGHFLRGY 427
QY 477 DLDLCLSSANSLYDDI 492
Db 428 DLPEHISNPEDVHRSI 443
RESULT 10
AAW38426
ID AAW38426 standard; Protein; 451 AA.
XX
AC AAW38426;
XX
DT 08-JUN-1998 (first entry)
XX
DE Human multiple myeloma oncogene 1 (MUM-1) product.
XX
KW Multiple myeloma oncogene 1; MUM1 protein; tumour; human;
KW diagnosis; interferon regulatory factor 4; IRF4.
XX
OS Homo sapiens.
XX
PN WO9745106-A1.
XX
PD 04-DEC-1997.
XX

the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. CC Polypeptides with growth factor activity may be used in cell cultures to CC promote cell growth. For example, such polypeptides may be used to CC manipulate stem cells in culture to give rise to neuroepithelial cells CC that can be used to augment or replace cells damaged by illness, CC autoimmune disease or accidental damage. The polypeptides and nucleotides CC may also be used in the diagnosis of the above conditions, and in drug CC screening techniques. The present sequence represents a novel human CC polypeptide of the invention.

XX SQ Sequence 72 AA;

Query Match 12.3%; Score 335.5; DB 22; Length 72;
Best Local Similarity 90.1%; Pred. No. 4.4e-18;
Matches 64; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 152 GGPPGFLATHAGLQAPGLPAPAGDKGDLQLLQVQSCDLHLTASWGA-DPVPTKA 210

Db 1 GGPPGFLATHAGLQAPGLPAPAGDEGDLALLAVQSCDLHLTASWGKDIPTKA 60

QY 211 PGGGQGLPLT 221

Db 61 LGEGQEGPLT 71

RESULT 12

ABBI1963
ID ABBI1963 standard; peptide; 392 AA.

XX AC ABBI1963;

XX DT 11-JAN-2002 (first entry)

XX DE Human IFN regulatory factor homologue, SEQ ID NO:2333.

XX KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer.

XX OS Homo sapiens.

XX PN WO200157188-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US03800.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-457740/49.

DR N-PSDB; ABA09207.

XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -

XX Claim 20; Page 287-288; 1963pp; English.

XX Sequences ABB10981-ABBI12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoietic regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

XX SQ Sequence 392 AA;

Query Match 11.7%; Score 320; DB 22; Length 392;

Best Local Similarity 26.0%; Pred. No. 4.8e-16;

Matches 119; Conservative 51; Mismatches 176; Indels 112; Gaps 19;

QY 62 AWAVARGWPPSSRGGPPPEAETRAERAGWKNFRCALRSTRFVLMRDNSSGDPADPHKV 121

Db 15 AWAEATGAYVP---GRDKPLPT-----WKNFRSALNRKEGLRLAEDRSKDPDHPHKI 65

QY 122 YALSRELWCWREGPGTDQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLPAPAGDKGD 181

Db 66 YEFV-----NSGVG-----DFSQDTSPTNNG-----GSTSD 93

QY 182 LLLQAVQSCDLADHLITASWGDAPVPTKAPGEGQEGPLPTGACAGGPGLPAGELYGWAVE 241

Db 94 -----TQEDILDELL-GNVLAPLP-----DGPFP-----SLA 120

QY 242 TTTPSPGPQPAALTGTEAAAPESPFOAEPYLSPSPACTAQVPE-SPGALDVTIMYGRTV 300

Db 121 VAPEPCQPQLRSPSLDNPTF-----FPNLGSEPNLKLKLLVPGEEWEFEVTAFYGRQV 174

QY 301 LOKVVGHSCTFLYGGPPDPVAVRATDPQQAFFSPA-ELPDQKQLRYTEELLRHVAPGLHL 359

Db 175 FOOTISCPGGLRVG-SEVGDRTLPGWPTLPDPGMSLTDRCGWSYVRHVLSCIGGLAL 233


```
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.

PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057669.
PR 05-SEP-1997; 97US-0057761.
PR 12-SEP-1997; 97US-0058785.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
XX WPI; 1998-506364/43.
DR N-PSDB; AAV59565.
DR
XX
XX New isolated human genes and the secreted polypeptide(s) they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
XX Claim 1; Page 573-574; 721pp; English.
XX
XX This sequence represents a secreted human protein encoded by the nucleic
CC acid molecule designated Gene 54 from the human cDNA clone HPRCU93
CC (deposited as clone ATCC 97899 and ATCC 209045).
CC The gene can be used to generate fusion proteins by linking to the gene
CC to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the
CC stability of the fused protein as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: AAV59511-V59812; amino acid sequences AAW4731-W5026)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 186
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAV59511 for described uses).
XX
XX Sequence 373 AA;

Query Match 9.5%; Score 259.5; DB 19; Length 373;
Best Local Similarity 25.2%; Pred. No. 1.6e-11;
Matches 100; Conservative 59; Mismatches 165; Indels 73; Gaps 15;

QY 115 PADPHKVVYALSRELWCREGPGTDTAEAPAAVPPQGGFFLAHTHAGLQAPGLPA 174
Db 9 PMNPVKIY----QVC-----DIPQPG-----SIINPGSTGS 36
QY 175 PAGDKGDLILLQVQSQSLADHLITASWGADPVTATPAGSGQGLPL-----TGACAGGPG 230
Db 37 APWDEKD---NDVDEDEDEDE-LDQSQHHVPTQDTFFFLNINGSMPAPASVGNCS 92
QY 231 PAGELYGWAVETTPSPGPQPAALTGGAAPSPHOAEPLYSPSPACTAVQSPGALD 290
Db 93 PEAV---W-----PKTEPLEMEVQ---APIQPFYSSPELWISSLPWT-----DLD 132
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:46:08 ; Search time 27 Seconds
(without alignments)
548.138 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731

Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSLYDDIECFMLELEQPA 503

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2695	98.7	503	US-08-999-774A-10	Sequence 10, Appl
2	454	16.6	427	US-08-705-771-16	Sequence 16, Appl
3	376	13.8	450	US-08-611-280-2	Sequence 2, Appl
4	376	13.8	450	US-09-195-940-2	Sequence 2, Appl
5	376	13.8	450	US-09-562-466-2	Sequence 2, Appl
6	361.5	13.2	450	US-08-611-280-25	Sequence 25, Appl
7	361.5	13.2	450	US-09-195-940-25	Sequence 25, Appl
8	361.5	13.2	450	US-09-562-466-25	Sequence 25, Appl
9	357	13.1	451	US-08-654-482-14	Sequence 14, Appl
10	328.5	12.0	393	US-09-230-371A-29	Sequence 29, Appl
11	306	11.2	425	US-09-230-371A-30	Sequence 30, Appl
12	259.5	9.5	373	US-09-149-476-374	Sequence 374, App
13	228	8.3	107	US-08-654-482-5	Sequence 5, Appl
14	222.5	8.1	108	US-08-654-482-1	Sequence 1, Appl
15	222.5	8.1	108	US-08-654-482-2	Sequence 2, Appl
16	218	8.0	107	US-08-654-482-7	Sequence 7, Appl
17	212	7.8	107	US-08-654-482-6	Sequence 6, Appl
18	166	6.1	108	US-08-654-482-4	Sequence 4, Appl
19	161.5	5.9	108	US-08-654-482-3	Sequence 3, Appl
20	155	5.7	561	US-08-842-255-52	Sequence 52, Appl
21	154.5	5.7	1064	US-08-642-255-62	Sequence 62, Appl
22	153.5	5.6	960	US-09-219-849-5	Sequence 5, Appl
23	152	5.6	355	US-08-483-533-41	Sequence 41, Appl
24	152	5.6	355	US-09-283-471A-41	Sequence 41, Appl
25	152	5.6	355	PCT-US91-06532-3	Sequence 3, Appl
26	148.5	5.4	560	US-07-683-957B-1	Sequence 1, Appl
27	148	5.4	252	US-08-642-255-61	Sequence 61, Appl

28	148	5.4	696	3	US-08-906-865-4	Sequence 4, Appl
29	148	5.4	696	4	US-09-129-868-4	Sequence 4, Appl
30	148	5.4	1057	3	US-08-931-820-4	Sequence 4, Appl
31	148	5.4	1461	4	US-09-585-887-9	Sequence 9, Appl
32	148	5.4	1461	4	US-09-289-578-9	Sequence 9, Appl
33	146	5.3	504	4	US-09-219-849-3	Sequence 3, Appl
34	145	5.3	1384	4	US-08-976-255-11	Sequence 11, Appl
35	144	5.3	777	1	US-08-642-255-53	Sequence 53, Appl
36	143.5	5.3	1057	3	US-08-931-820-1	Sequence 1, Appl
37	143.5	5.3	1341	3	US-08-963-825-18	Sequence 18, Appl
38	143.5	5.3	1341	4	US-09-500-811-18	Sequence 18, Appl
39	143.5	5.3	1341	4	US-09-570-573-18	Sequence 18, Appl
40	143.5	5.3	1341	4	US-09-548-608-18	Sequence 18, Appl
41	140.5	5.1	234	1	US-07-609-716-66	Sequence 51, Appl
42	140	5.1	357	1	US-08-642-255-33	Sequence 66, Appl
43	140	5.1	357	4	US-08-475-411A-66	Sequence 66, Appl
44	140	5.1	357	4	US-08-478-029A-66	Sequence 66, Appl
45	140	5.1	357	4		

ALIGNMENTS

RESULT 1
US-08-999-774A-10
; Sequence 10, Application US/08999774A
; Patent No. 6274312
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Seghezzi, Wolfgang
; APPLICANT: Shanahan, Frances
; APPLICANT: Lees, Emma M.
; APPLICANT: McClanahan, Terrill K.
; TITLE OF INVENTION: Intracellular Regulatory Molecules;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: PatentIn Release #1.0, Version #1.30
FILING DATE: 10-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,818
FILING DATE: 11-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0646
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-774A-10

Query Match 98.7%; Score 2695; DB 4; Length 503;
Best Local Similarity 99.0%; Pred.No. 3e-223; 5; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 5;

```
QY 1 MALAPERAARVLFGEWLLGEISGCVYEGWLDDEARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAARVLFGEWLLGEISGCVYEGWLDDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGRPPSSRGSGPPPEAEATRAAGWKNFRCAIRSTRFRFVMLRDNSSGDPADPHK 120
DB 61 KAWAVARGRPPSSRGSGPPPEAEATRAAGWKNFRCAIRSTRFRFVMLRDNSSGDPADPHK 120
QY 121 VYALSRELWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLPAPAGDKG 180
DB 121 VYALSRELWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLPAPAGDKG 180
QY 181 DILLQAVQOCLADHLLTASWAGDPVPTKAPGEGQGLPLTGACAGGPGPLPAGELYGWAV 240
DB 181 DILLQAVQOCLADHLLTASWAGDPVPTKAPGEGQGLPLTGACAGGPGPLPAGELYGWAV 240
QY 241 ETTSPGQPAALTTGAAAPESPQAEPLSPSPSACTAVQPSFGALDVTIMYKRTV 300
DB 241 ETTSPGQPAALTTGAAAPESPQAEPLSPSPSACTAVQPSFGALDVTIMYKRTV 300
QY 301 LQKVGHPSCTELYGPPDPRAVRATDPOVAFPSPAELPDQKQLRYTEELLRHVAPGLHLE 360
DB 301 LQKVGHPSCTELYGPPDPRAVRATDPOVAFPSPAELPDQKQLRYTEELLRHVAPGLHLE 360
QY 361 LRGPQWARMGKCKVYWEVGGPPGSGASPTPACLLPRNCDTPIFDPRVFQELVEFRAR 420
DB 361 LRGPQWARMGKCKVYWEVGGPPGSGASPTPACLLPRNCDTPIFDPRVFQELVEFRAR 420
QY 421 QRRGSPRYTYLFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSSDL 480
DB 421 QRRGSPRYTYLFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSSDL 480
QY 481 CLSSANSYDDIECFLEMQPA 503
DB 481 CLSSANSYDDIECFLEMQPA 503

RESULT 2
US-08-705-771-16
; Sequence 16, Application US/08705771
; Patent No. 6054289
; GENERAL INFORMATION:
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
; APPLICANT: Jian Ni and Jing-Shan Hu
; TITLE OF INVENTION: Human Genes, Sequences and
; TITLE OF INVENTION: Expression Products
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,771
; FILING DATE: August 30, 1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-346 (PF196)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 427 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-705-771-16

Query Match 16.6%; Score 454; DB 3; Length 427;
Best Local Similarity 28.0%; Pred. No. 4.8e-31;
Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;

QY 10 PRVLFGEWLLGEISGCVYEGWLDDEARTCFRVPWKHFARKDLSEADARIFKAWAVARG 69
DB 6 PRXL--PWLVSQDLGLEGVAMWVKSRTFRFVMLRDNSSGDPADPHK 63
QY 70 WPSSRGSGPPPEAEATRAAGWKNFRCAIRSTRFRFVMLRDNSSGDPADPHK 129
DB 64 YVP---GRDKPLPT---WKRNFRSALNRKEGLRLAEKDRSKDPHDKHIFEFV--- 110
QY 130 WREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLPAPAGDKGLLQAVQ 189
DB 111 -NSGVG---DFSQPTSDTNGG-----GSTSD-----TQ 136
QY 190 SCLADHLLTASWAGDPVPTKAPGEGQGLPLTGACAGGPGPLPAGELYGWAVETTPSPGP 249
DB 137 EDILDELL--GNMVLAPLP-----DGPDP-----SLAVAPEPCPQ 169
QY 250 PAALTGEEAAPPSPHQAEPYLSPPSACTAVQPSFGALDVTIMYKRTVQKVVGHP 308
DB 170 PLRSPSLDNTPT-----FPNLGSPENPLKRLLVGEEWEFEVTAFYRGVQVFOQTISCP 223
QY 309 SCTFLYGPDPDAVRATDPOVAFPSPA-ELPDQKQLRYTEELLRHVAPGLHLELRGPQLW 367
DB 224 EGLRVG-SEVGDRTLPCHPVTLPDFGMSLTDRGVMSYRVHVLSCGGLALWRAGOWLN 282
QY 368 ARRMGKCKVYWEVG-----GPPGSGASPTPACLLPRNCDTPIFDPRVFQELVEFR 418
DB 283 AQLRGHCHTYWAVSEELLPNSGHGPDE-----VPKDEGGVFDLGPFIVDLITFT 333
QY 419 ARQRRGSPRYTYLFGQDLSAGRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSSDL 478
DB 334 EGSGR-SPRYALMFCVGSWPDQDPWTKRLVMVYVPTCLRALVEMARVGGASSLENT-V 391
QY 479 DCLSSANSLSL---YDDIECFLEML 499
DB 392 DLHISNSHPLSLTSDQYKAYLQDL 415

RESULT 3
US-08-611-280-2
; Sequence 2, Application US/08611280
; Patent No. 5891666
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Grossman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,280
; FILING DATE:
```

```
/
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Oleski, Nancy A.
/ REGISTRATION NUMBER: 34,688
/ REFERENCE/DOCKET NUMBER: A-338A
/ INFORMATION FOR SEQ ID NO. 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 450 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
US-08-611-280-2

Query Match      13.8%; Score 376; DB 2; Length 450;
Best Local Similarity 28.1%; Pred. No. 2.5e-24;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

QY 16 EWLGEISSGCVGELQWLDDEARTCFRVWKHPARKDLS-EADARIFKAWAVARGRWPPSS 74
Db 26 QWLIDQIDSGKYPGLVWNEEKSVFRIPWKHAGKQDYNREDAALFKAWALFKGKF---- 81
QY 75 RGGGPPPAETAERAGWKNFRCALRSTRFVMLRDNAG-DPADPHKVVYALSRELWCWREG 133
Db 82 REGIDKDPPT-----WKTRLCALNKSNDFEELVERSOLDISDPKYVTRI----- 127
QY 134 PGTDQTEAEPAAVPPPGGPPGP---FLAHTHAGLOAPGPPAPAGDKGDLILLQAVQOS 190
Db 128 -----VPEGAKKGAKQLTLDQTMAMGHYPMTAP---YGSPLPAQOVHNY 169
QY 191 CLADHLLTASGADVPYTKAPGEGEGLPLTGACAG-----GGLP-----AGELYGWAVE 241
Db 227 ESQAPGPIEPSIRSAEALA-----LSDC-----RLHICLYRDLIV 263
QY 301 LQKVVGHP-SCFTLYGPPDPVAVRATDQQVAFPPSPAEPLPDQKQLRYTBELLRHVAPGLHL 359
Db 264 KELTTTSPGCRISHG---HTVDVSNLDQVLPFP---DDNGQRKNIEKLLSHLERGLVL 317
QY 418 RARQRRGSPRYTILGFGODLSAGRPKEKSLVLVLEPWLCR 459
Db 372 -AHGGRPAFRFQVTLFCFGEFFDPDQ-RORKLITAHVEPLLAR 411

RESULT 4
US-09-195-940-2
/ Sequence 2, Application US/09195940
/ Patent No. 6258935
/ GENERAL INFORMATION:
/ APPLICANT: Matsuyama, Toshifumi
/ APPLICANT: Grossman, Alex
/ APPLICANT: Richardson, Christopher D.
/ TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Amgen Canada Inc.
/ STREET: 6733 Mississauga Road, Suite 303
/ CITY: Mississauga
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: L5N 6J8
/ COMPUTER READABLE FORM:
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30

/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/195,940
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/611,280
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Oleski, Nancy A.
/ REGISTRATION NUMBER: 34,688
/ REFERENCE/DOCKET NUMBER: A-338A
/ INFORMATION FOR SEQ ID NO. 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 450 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
US-09-195-940-2

Query Match      13.8%; Score 376; DB 4; Length 450;
Best Local Similarity 28.1%; Pred. No. 2.5e-24;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

QY 16 EWLGEISSGCVGELQWLDDEARTCFRVWKHPARKDLS-EADARIFKAWAVARGRWPPSS 74
Db 26 QWLIDQIDSGKYPGLVWNEEKSVFRIPWKHAGKQDYNREDAALFKAWALFKGKF---- 81
QY 75 RGGGPPPAETAERAGWKNFRCALRSTRFVMLRDNAG-DPADPHKVVYALSRELWCWREG 133
Db 82 REGIDKDPPT-----WKTRLCALNKSNDFEELVERSOLDISDPKYVTRI----- 127
QY 134 PGTDQTEAEPAAVPPPGGPPGP---FLAHTHAGLOAPGPPAPAGDKGDLILLQAVQOS 190
Db 128 -----VPEGAKKGAKQLTLDQTMAMGHYPMTAP---YGSPLPAQOVHNY 169
QY 191 CLADHLLTASGADVPYTKAPGEGEGLPLTGACAG-----GGLP-----AGELYGWAVE 241
Db 170 MPPHP--DRSW-RDYAPDQSHPEIPYQCFVTGPRGHHWQSPSCNGCQVTFYACAPP 226
QY 242 TTPSPG-PQPAALTGEEAAAPESPHQAEPLSPSACTAVQEPSFGALDVTIMYKGRV 300
Db 227 ESQAPGPIEPSIRSAEALA-----LSDC-----RLHICLYRDLIV 263
QY 301 LQKVVGHP-SCFTLYGPPDPVAVRATDQQVAFPPSPAEPLPDQKQLRYTBELLRHVAPGLHL 359
Db 264 KELTTTSPGCRISHG---HTVDVSNLDQVLPFP---DDNGQRKNIEKLLSHLERGLVL 317
QY 360 ELRGQPLWARRMGKCKVWVEGGPPGSGAS--PSTPACLLPRNCDTPIPDFRVFFQELVEF 417
Db 318 WMAPDGLYAKRLCQSRITYD--GPLALCSDRPNK----LERDQCKLFDTQOQFLSELOVF 371
QY 418 RARQRRGSPRYTILGFGODLSAGRPKEKSLVLVLEPWLCR 459
Db 372 -AHGGRPAFRFQVTLFCFGEFFDPDQ-RORKLITAHVEPLLAR 411

RESULT 5
US-09-562-466-2
/ Sequence 2, Application US/09562466
/ Patent No. 6369202
/ GENERAL INFORMATION:
/ APPLICANT: Matsuyama, Toshifumi
/ APPLICANT: Grossman, Alex
/ APPLICANT: Richardson, Christopher D.
/ TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Amgen Canada Inc.
/ STREET: 6733 Mississauga Road, Suite 303
/ CITY: Mississauga
/ STATE: Ontario
/ COUNTRY: Canada
```


Db 225 PESQAGVPTPEPSIRSAEALA-----FSDC-----RLHICLYREIL 261
QY 300 VLQKVVGHP--SCTFLYGGPPDPAVRATDPOQVAFPSPAELPDQKQLRYTELLRHVAPGLH 358
Db 262 VKSLTSSPGCRISHG---HTYDASLMDQVLPYP---EDNQKRNIEKLSHLRGVV 315
QY 359 LELRGPOLWARRMGKCKVYVWVGPPGSAS--FSTPACLLPRNCDTPIFDPRVFFOELVE 416
Db 316 LWMADGLYAKRLCQSIYWD--GPLALCNDPNK---LERDQCKLFDTCQFLSELQA 369
QY 417 FRARQRGSPRYTIYLGFGQDLGAGRPKESLVLVLEPWLCRVHLEGTORREGVSSLDSS 476
Db 370 F-AHGRSLRFRQVTLFCGEEFPDPQ--RQRKLITAHVEPLLARQLYFAQQNSGHLRGY 427
QY 477 DLDLCLSSANSLYDDI 492
Db 428 DLPEHISNPEDYHRSI 443
RESULT 9
US-08-654-482-14
; Sequence 14, Application US/08654482
; Patent No. 6245562
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; TITLE OF INVENTION: IDENTIFICATION OF GENES ALTERED IN
; TITLE OF INVENTION: MULTIPLE MYELOMA
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/654,482
; FILING DATE: 28-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 50995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-654-482-14
Query Match 13.1%; Score 357; DB 4; Length 451;
Best Local Similarity 27.3%; Pred. No. 1.1e-22;
Matches 135; Conservative 73; Mismatches 194; Indels 92; Gaps 24;
QY 16 EMLGEISSCGYEGLOWLDEARTCFRVPWKHFARKDLS--EADARIFKAWAVARGWPPSS 74
Db 26 QWLIDQIDSGKYPGLWENEEKSIFRIPWKHAGQDYNBEDAALFKAWALFKGK---- 81
QY 75 RGGGPPPEAETAERAGWKTNFRCALRSTRFRVLMRLNSG--DPADPHKVALSRELWCWREG 133
Db 82 REGIDKEDPT-----WKTRLCALNKSNDFEELVERSQLDISDPKYKVIIVEGAKK-- 134
QY 134 PGTDQTEAEAPAAVPPPGGPPGPFLLAHTHAGLAQPGP-LPAPAGDKGDLILLQAVQOSCL 192

Db 135 -GAKQLTLEDPOQ-----SMSPYT-MTTPYPSLPA-----QQVHNYMM 171
QY 193 ADHLTASAGADVPPTKAPGEGQGLPT-----GACAGGGLPAGELYGWAVET 242
Db 172 PP--LDRSW-RDVPDPHPPEIPYQCPMTFGPRGHWQGPACENGCOV-TGTIYACAPPE 227
QY 243 TPSPG-PQPAALTTGSAAPESPHQAEPLSPSPSACTAVQEPSPGALDVTIMYKGRTVL 301
Db 228 SQAGVPTPEPSIRSAEALA-----FSDC-----RLHICLYREILVK 264
QY 302 QKVVGHP--SCTFLYGGPPDPAVRATDPOQVAFPSPAELPDQKQLRYTELLRHVAPGLHLE 360
Db 265 ELTTSPGECRISHG---HTYDASLMDQVLPYPEDNGHRKNI---ENLLSHLERGVVLW 318
QY 361 LRGPOLWARRMGKCKVYVWVGPPGSAS--PSTPACLLPRNCDTPIFDPRVFFOELVEFR 418
Db 319 MAPDGLYAKRLCQSIYWD--GPLALCNDPNK---LERDQCKLFDTCQFLSELQA- 371
QY 419 ARQRGSPRYTIYLGFGQDLGAGRPKESLVLVLEPWLCRVHLEGTORREGVSSLDSSDL 478
Db 372 AHGRSLRFRQVTLFCGEEFPDPQ--RQRKLITAHVEPLLARQLYFAQQNSGHLRGYDL 430
QY 479 DLCLSSANSLYDDI 492
Db 431 PEHISNPEDYHRSI 444
RESULT 10
US-09-230-371A-29
; Sequence 29, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Human
US-09-230-371A-29
Query Match 12.0%; Score 328.5; DB 4; Length 393;
Best Local Similarity 25.7%; Pred. No. 2.5e-20;
Matches 120; Conservative 69; Mismatches 179; Indels 99; Gaps 19;
QY 17 WLGEISSCGYEGLOWLDEARTCFRVPWKHFARKDLS--ADARIFKAWAVARGWPPSSR 75
Db 15 WVVEQVESGQFPVCVCHDDTAKTMRIPWKHAGQDREODAAFFKAWALFKGKY----- 69
QY 76 GGGPPPEAETAERAGWKTNFRCALRSTRFRVLMRLNSG--DPADPHKVALSRELWCWREGP 134
Db 70 -----KEGDTGGPAVWKTRLCALNKSSEPFKEVPERGRMDVAEPKYVQLLPPGIVSGQP 124
QY 135 GTDQTEAEAPAAVPPPGGPPGPFLLAHTHAGLAQPGP-LPAPAGDKGDLILLQAVQOSCLAD 194
Db 125 GTQKVPK-----ROHSSVS-----SERKEED-----AMQNCITLSP 156
QY 195 HLITASAGADVPPTKAPGEGQGLPLTGACAGGGLPAGELYGWAVETTPSPGPQPAALT 254
Db 157 SVLQDSLNNEE-----EGASG---GAVHSDIG-----SSSSSSSPPEQVETDT 196
QY 255 TGEAAPAPESPHQAE-----PYLSPSPSACTAVQEPSPG-ALDVTIMYKGRTVLQKVGHPS 309

Db 197 T-----EAPFGDQSRSLFLP-----PEPDYSLITFIYNGRVGAEQVQSLD 240
Qy 310 CTFLYGPDPVAVRATDPOQVAPSPAPLDPKQRLRYTEBELLRHVAPGLHLELRGPOLMAR 369
Db 241 CRLV---AEPGSESSMEQVLPKPGLEP-----TQRLLSQLERGILVNASPGLFVQ 291
Qy 370 RMGKCKVYWEVGGPGSASPSPACLLPNCDDTPIPIDFRVFQELVEFRARQRRG-SPRY 428
Db 292 RLCPPIISWNAFQAPPGPH-----LLPSNECVLEFRTAYFCRDLY--RYFQGLGPPPKF 345
Qy 429 TTVLFGQDLSAGRPKEKSLVLKLPWLCRVHLEGT--OREGVSSL 473
Db 346 QVTLNFWESHGSSHTPQNLITVQMEQAFARYLLEQTPEQQAAILSL 392

RESULT 11
US-09-230-371A-30
; Sequence 30, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230, 371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 30
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Human
US-09-230-371A-30

Query Match 11-2%; Score 306; DB 4; Length 425;
Best Local Similarity 23.5%; Pred. No. 2.4e-18;
Matches 123; Conservative 66; Mismatches 168; Indels 166; Gaps 22;

Qy 16 EWLGEISSGCGEGLQWLDEARTCFVPWKHFARKDLS-EADARIPKAWAVARGRWPPSS 74
Db 12 QWLEIQIDSSMYPLIWEENEKSMFRIPKHAGKQDYNQVDASIFKAWAVFKGFE--- 67
Qy 75 RGGGPPPEAETFRAGWKNFRCAURSTRFRVWLRNDSG-DPADPHKVYALSREL----- 128
Db 68 -----KEGDKAEPATWTKRLCALNKSDFBEVTRDSQLDISSEPKYVRIVPEEDQCK 121
Qy 129 -----CWREGGTDTEAEAPAAVPPQGGPGFLAHTHAGLQAPGLPAPAGDKGD 181
Db 122 LGVATAGCVNEVTEMECGSEIDELKEPS-----VDDYGMGIMKRSPPDA----- 168
Qy 182 LLLQAVQOQSLADHLITAGSAGDPPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVE 241
Db 169 -----CRSQ--LLPDM-----WAHE 181
Qy 242 TTPSPGPQPAALTGTGAAPAPESPHQAEPLYSPSPSACTAVQSPSGALDVTIMYKGRVYL 301
Db 182 --PSTGRR--LVTGYTTY-DAHSAF-----SQMWISFYVGGKLVG 217
Qy 302 QKVGHGHP-SC-----TFLYCPDPAVAVRATDPOQVAPSPAPLDPKQRLYTEEL 349
Db 218 QATTTCPCECRISLSQPGLPGLTKLYG-----EGLELVRFPPADTIPESEQRQVTRKL 270
Qy 350 LRHVAPG-LHLELRGPOLWARMGCKVYWEV-----GGPPGSASBSTACLLPNCDDT 403
Db 271 FGHLERGVLSHRSQG--VFVKLCQGRVFCVVVVVCKGRPNK-----LERDEVVQ 319

Qy 404 IFDPRVFQELVEFRARQRRGSPRYTIYLGFGQDLSAGRPKEKSLVLKLPWLCRVHLE 463
Db 320 VEDTSQIFRELQOQFYNSQGR-LPDGRVVLFCFGEFFPDNAPRSKILVQIE-----QLYVR 374
Qy 464 GTOREGVSSLDSSDL-----DLCSSANSLY 489
Db 375 QLAEEAGKSCGAGSVQWQAPDEEPPDQVFRMFPDICASHORSFF 417

RESULT 12
US-09-149-476-374
; Sequence 374, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149, 476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23


```

; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/654,482
; FILING DATE: 28-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 50995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-654-482-2

```

```

Query Match      8.1%; Score 222.5; DB 4; Length 108;
Best Local Similarity 41.2%; Pred. No. 5.6e-12;
Matches 47; Conservative 21; Mismatches 35; Indels 11; Gaps 4;

Qy      16 EMLIGEISSGCEGLQWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPPSS 74
Db      4 QWLIDQIDSGKYPGLVWNEEKSVFRIIPWKHAGQDYNREEDAALFKAWALFKGF---- 59

Qy      75 RGGGPPPEAETAEAGWKTFRALSTRFRVLMRDNSG-DPADPHKVVALSRE 127
Db      60 REGIDKDPPT-----WKTRLCALNKNSDFEELVERSOLDISDPYKYVIVPE 108

```

Search completed: June 18, 2003, 12:48:51
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:47:53 ; Search time 50 Seconds
(without alignments)
1088.561 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSYDDIEFLMELEQPA 503

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/FCI_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/FCIUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	454	16.6	427	9	US-09-975-253-2
2	259.5	9.5	373	9	US-09-809-391-374
3	167	6.1	1023	9	US-09-893-519A-14
4	153	5.6	349	10	US-09-919-497-76
5	150.5	5.5	550	9	US-09-976-740-47
6	150.5	5.5	550	12	US-10-023-529-47
7	150.5	5.5	550	12	US-10-023-523-47
8	148.5	5.4	560	9	US-09-813-398-23
9	148.5	5.4	560	9	US-10-122-805-4
10	148	5.4	696	9	US-10-060-036-159
11	148	5.4	1464	9	US-10-171-311-36
12	148	5.4	1464	9	US-10-216-705-21
13	148	5.4	1464	9	US-10-149-352-2
14	148	5.4	1464	9	US-10-060-036-166
15	145.5	5.3	1516	9	US-10-157-031-128
16	143	5.2	362	9	US-10-050-882-68
17	142	5.2	394	9	US-10-050-882-137
18	142	5.2	394	9	US-10-260-715-8
19	140.5	5.1	1259	9	US-10-260-715-8

20	139	5.1	1207	10	US-09-927-112-2	Sequence 2, Appli
21	138	5.1	325	10	US-09-927-112-10	Sequence 10, Appl
22	138	5.1	376	10	US-09-955-518-20	Sequence 20, Appl
23	137	5.0	538	9	US-09-976-740-43	Sequence 43, Appl
24	137	5.0	538	12	US-10-023-529-43	Sequence 43, Appl
25	137	5.0	538	12	US-10-023-523-43	Sequence 43, Appl
26	137	5.0	714	9	US-10-233-885-44	Sequence 44, Appl
27	137	5.0	714	9	US-10-231-581-44	Sequence 44, Appl
28	137	5.0	897	9	US-10-099-895-1	Sequence 1, Appli
29	136.5	5.0	527	9	US-09-854-133-216	Sequence 216, App
30	136.5	5.0	527	10	US-09-738-973-216	Sequence 216, App
31	136	5.0	638	9	US-10-038-010-4	Sequence 4, Appli
32	134	4.9	503	9	US-10-078-547-2	Sequence 2, Appli
33	134	4.9	591	9	US-09-291-417-103	Sequence 103, App
34	134	4.9	699	9	US-10-121-988-143	Sequence 143, App
35	133.5	4.9	731	9	US-10-086-464-17	Sequence 17, Appl
36	133.5	4.9	1274	9	US-10-020-215-2	Sequence 2, Appli
37	133	4.9	881	10	US-09-816-860A-2	Sequence 2, Appli
38	132.5	4.9	417	10	US-09-963-285-4	Sequence 4, Appli
39	132.5	4.9	501	10	US-09-963-285-2	Sequence 2, Appli
40	132.5	4.9	501	10	US-09-963-285-10	Sequence 10, Appl
41	132.5	4.9	507	9	US-10-078-547-24	Sequence 24, Appl
42	132	4.8	1647	9	US-09-824-574-4	Sequence 4, Appli
43	131.5	4.8	945	8	US-08-965-272-2	Sequence 2, Appli
44	131.5	4.8	945	12	US-10-121-882-2	Sequence 2, Appli
45	131.5	4.8	1339	9	US-10-082-830-282	Sequence 282, App

ALIGNMENTS

RESULT 1
US-09-975-253-2
; Sequence 2, Application US/09975253
; Patent No. US20020164694A1
; GENERAL INFORMATION:
; APPLICANT: Paul Moore et al.
; TITLE OF INVENTION: Methods and Compositions for Treating and Preventing Infection
; FILE REFERENCE: PF196P1
; CURRENT APPLICATION NUMBER: US/09/975,253
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/239,963
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals Ile, Leu, Phe, or Val

Query Match	16.6%	Score 454;	DB 9;	Length 427;
Best Local Similarity	28.0%	Pred. No. 2.1e-21;		
Matches 141;	Conservative 63;	Mismatches 192;	Indels 108;	Gaps 19;
QY	10	PRVLFGWLLGISGCGYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVGR	69	
Db	6	PRXL--PWLVSQDLQLEGVAWVANKSRTRFRIPWKHGLRQDAQQQDFGIFQAWAETGA	63	
QY	70	WPSSRGCGPPPEAETAEAGKHTFRCALRSTRFRVLMRLDNGSDPADPHKVYALSRELC	129	
Db	64	YVP---GRKDPDLPT-----WKNFRSALNKKEGLRLAEDRSKDPDPHKYIEFV----	110	
QY	130	WREGPGTQDTEAEAPAAVPPQCGPPGPPFLAHTAGLQAPGFLPAPAGDKGDLLOAVQ	189	
Db	111	-NSGVG----DFSQPDTSPTDNGG-----GSTD-----TQ	136	
QY	190	SCLADHLLTASWAGADVPVTKAPGEGQEGPLTGCACAGGGLPAGELYGNWAVETTPSPGPQ	249	

```
Db 137 EDILDELL-GNMVLAPLP-----DPGPP-----SLAVAPPCPQ 169
QY 250 PAALTGCEAAAPSPHOAEPYLSPPSACTAVQEP-SPGALDVTIMYKGTVLQKVGH 308
Db 170 PLRSPSLDNTP-----FNLGSENPRLKRLILVPGSEWFEFVTAFYRGQVFOQTISC 223
QY 309 SCTFLYGPDPAPVRATDPQOQVAPSPA-ELPDQKQLRYTEELRHVAPGLHLELPGQLW 367
Db 224 EGLRLVG-SEVGDTLPGWPTLPDQMSLTDGVMYSYVRHVLSCGGGLALWRAGOWLW 282
QY 368 ARMKGKVKVWEVG-----GPGSASPTPACLLPRNCPTPIPDFRVFQELVEFR 418
Db 283 AQLRGHCHTYWAVSEELLNPSGHGPDGE-----VPKDEKGGVFDLPGFVIDLTFT 333
QY 419 AQORRGSPRYTILYFGQDLSAGRPKEKSLAVKLEPWLCRVHLEGTQREGVSSLOSDL 478
Db 334 EGSGR-SPRYALMFCVGESWPQDPMTKRLVMKVVPTCLRALVEMARVGGASSENT-V 391
QY 479 DCLSSANSI---YDDIECFMEL 499
Db 392 DLHISNHPLSLTSQYKAYLODL 415
```

RESULT 2

```
US-09-809-391-374
; Sequence 374, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002p2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 374
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (373)
; OTHER INFORMATION: Xaa equals stop translation
US-09-809-391-374
```

```
Query Match 9.5%; Score 259.5; DB 9; Length 373;
Best Local Similarity 25.2%; Pred. No. 3.7e-09;
Matches 100; Conservative 59; Mismatches 165; Indels 73; Gaps 15;

QY 115 PADPHKLYALSRELWCWREGTQDTEAEAPAAVPPQGGPPGPFPLAHTHAGLQAPGLPA 174
Db 9 PMNPVKIY---QVC-----DIPQPG-----SIINPGSTGS 36

QY 175 PAGDKGDLLOAVQOQSLADHLLTASWGADPVPTKAPGEGQGLPL-----TGACAGPGL 230
Db 37 APWDEKD---NDVDEEDEE-LDQSQHHVPIQDTFFFLNINGSMPAPASVGNCSVGNCS 92

QY 231 PAGELYGWAVETTPSPGQPAALTGTGAAPESPHOAEPLYLSPSACTAVQSPGALD 290
Db 93 PEAV---W-----PTELEMEVPO--APIQFFYSSELWISSLPMT-----DL 132

QY 291 VTIYKGTVLQKV-VGHP-SCITFLYGPDPFAVRATD-----PQVAFPSAPLDPQKQ 342
Db 133 IKQYRGKEYGQMTVNSPNQGCRLFYGLDGMPPDQELFGVXLEQVKFGPGEHINEKQ 192

QY 343 LRTEELLRHVAPGLHLELPGQLWARMKCKYKVEVGGPPGSGASPSSTACILPNC 402
Db 193 KLFTSKLLDVMRGLILEVSGHAIYAIRLQCKQVYWSGGPCAPSLVAPN-----LIERQKV 248
```

```
QY 403 PIFDRVVFQELVEFRARORRGSPRYTILYFGQDLSAGRPKEKSLVVKLEPWLCRVHL 462
Db 249 KLFCLTFLSLDLIAHQKQIEKQPPFEIYLCFGEWPDGKPLERKLLILVQVIVVARMY 308
QY 463 EGTQREGVSLDSSDLDLCLSSANSYLDIECFMEL 499
Db 309 EMFSGDFTRSFSGSVRLQISTPD-IKDNIVAGLKQL 344
```

RESULT 3

```
US-09-893-519A-14
; Sequence 14, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BURMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Thamara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARNITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAQ, Tariq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/IG548-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Corresponds to SEQ ID NO: 87
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Human Genbank/CAA72189
; DATABASE ENTRY DATE: 1997-06-25
; RELEVANT RESIDUES: (1)..(1023)
US-09-893-519A-14
```

```
Query Match 6.1%; Score 167; DB 9; Length 1023;
Best Local Similarity 24.6%; Pred. No. 0.0077;
Matches 102; Conservative 22; Mismatches 131; Indels 160; Gaps 20;

QY 62 ANAVAGRWPSPSRGGGPPPEAETABRAGWKTNFRCLALRSTRFRVMLRDNNSGDPADPHKV 121
Db 68 AGAAGAGPAAPAGAPAAPEPPAGRA-----RPGGGGPPORP--- 105

QY 122 YALSRELWCWREGTQDQ---TEAEAPAAV---PPQG--GPPGPFPLAHTHAGLQAPGL 172
Db 106 -----GPPSRRELVPAGPAPPAAKURPPPEGAGACAPVPA---AAAVAGPE 151

QY 173 PAPAG---DKGDLLOAVQOQSLADHLLTASWGADPVPTKAPGEGQGLP-----LTGAC 224
Db 152 PAPAGAKPAGPAALAA-----RAGPGPGPGPGPGPGKPGAGPAGAAQTLLNSA 200

QY 225 A-----GGPG-----LPAGELYGWAVETTPSPGQPAALTGTGAAPES 264
Db 201 ALINSHHAAAPAVSLVNNNGPAALLPLPKAAPGTIVTQTPPFVG-----AAAPPAP 250
```

QY 265 HQAEPYLSPSACTAVQBPSCALDVTIMYKGRVTLQKVGHPSCFTFLYGPDP----- 318
Db 251 AAPSPPAAPAPAPAAAPPPPPA-----PATLAPPGHPA-----GPETAAPVP 296
QY 319 -----PAVRATDPOOV-----APSP-----AELPDQQLRYTEEL 349
Db 297 PPAQAQNGSAGAPAPAPAGAGVQCPGPGGAAAPAPAGVRAESP-----KKV 348
QY 350 LRHVAPGLH-LBLRGQLWARRMGKCKVYMEVGGPGSPSPACLLPRNCDDTP 403
Db 349 VQAAPPAATLAASGPASTAASM-----VIGPTMQALPSPAAPPAPGTP 395
RESULT 4
US-09-919-497-76
; Query Match 5.6%; Score 153; DB 10; Length 349;
; Best Local Similarity 34.0%; Pred. No. 0.019;
; Matches 35; Conservative 18; Mismatches 40; Indels 10; Gaps 3;
; ORGANISM: Homo sapiens
; US-09-919-497-76
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 76
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-919-497-76

Query Match 5.6%; Score 153; DB 10; Length 349;
Best Local Similarity 34.0%; Pred. No. 0.019;
Matches 35; Conservative 18; Mismatches 40; Indels 10; Gaps 3;
ORGANISM: Homo sapiens
US-09-919-497-76

QY 11 RVLFGEWLLGEISSGCGYGLQWLDEARTCFRVWKHFARKDLS-EADARIFKAWAVARGR 69
Db 5 RMRMRPLSEQINSNTIPGLKWLNEKKIFQIPWMAHARGHWDVEKADPLFRNRAIHTCK 64
QY 70 WPPSSRGCGPPPEAETAEARAGWKTNFRCLALRSTRFRVLMRDNS 112
Db 65 HOP-----GVDKDPKT-----WKANFRCAANSLPDIEEVKDKS 98

RESULT 5
US-09-976-740-47
; Sequence 47, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47

; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-976-740-47
Query Match 5.5%; Score 150.5; DB 9; Length 550;
Best Local Similarity 23.4%; Pred. No. 0.044;
Matches 82; Conservative 20; Mismatches 111; Indels 137; Gaps 15;
QY 2 ALAPERAAPRVLFGEWLLGEISSGCGYGLQWLDEARTCFRVWKHFARKDLSADARIFK 61
Db 20 AAAASSSAASPHYQEWILDTI-----DSLRS-----RKARPDLE-ERICRMVR 60
QY 62 AWAVARGWPPSSRCGGPPPEAETAE-----RAGWKTNFRCLALRSTRFRVLMRDNS 112
Db 61 -----RRHPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGA 107
QY 113 GDPADPHKVYALSRCLWREGPGTQTEAEAPAAVPPQGGPPGFLAHTAGLQAPGL 172
Db 108 TTPAPPRAP-----RGGP-----AAAAAPPTTAPPPP-----PAPV 139
QY 173 PAPAGDKGDLQLQAVQOSCLADHLLTASWGADPVPTKAPGEGGGLPTGCACAGSGPLPA 232
Db 140 AAAA-----APARAPRAAAAAAATAAPPSPGPAQPG 170
QY 233 GELYGAVETTPSPQCPQPAALTGTGAAPESPQHEPVLSPSPSACTAVQBPSCALDVT 292
Db 171 PR-----AQRAAPLAAPPAPAPAPAAAPPAGPRRA-----PPFAAAVAARE-SP----- 214
QY 293 IMYKGRVTLQKVGVHPSCTFLYGLPDPAVRATDPOQVAPFPPSPAEPLDQKQ 342
Db 215 -----LPPPPQ-----PAPPOQQQPPPPPPPPQ 241

RESULT 6
US-10-023-529-47
; Sequence 47, Application US/10023529
; Patent No. US20020129389A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-529-47
; Query Match 5.5%; Score 150.5; DB 12; Length 550;
; Best Local Similarity 23.4%; Pred. No. 0.044;
; Matches 82; Conservative 20; Mismatches 111; Indels 137; Gaps 15;
QY 2 ALAPERAAPRVLFGEWLLGEISSGCGYGLQWLDEARTCFRVWKHFARKDLSADARIFK 61

Db 20 AAAASSAASPHYQEWILDTI-----DSLRS-----RKARPD-ERICMVR 60
Qy 62 AWAVARGWPPSSRGSGPPPEAETA-----RAGWKTNFRCAL--RSTRFVMLRDNS 112
Db 61 -----RRHGPERTEAELEKLIQRAVLRSYKGSISYRNAARVQPPRGA 107
Qy 113 GDPADPHKVYALSRELWCWREGPGTDQTEAEAPAAVPPPGGPGPFLLAHTHAGLQAPG 172
Db 108 TPAPPAPR-----RGSP-----AAAAAAPPPTAPPPP-----PAPV 139
Qy 173 PAPAGDKDLLLLQAVQOSCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGGLPA 232
Db 140 AAAA-----APARAPRAAAAAAATAAPPSPGPAQPG 170
Qy 233 GELYGWAVETTPSPGPOPAALTTGEAAAPESPHQAEPLYLSPSPSACTAVQEPSGALDVT 292
Db 171 PR-----AQRAAPLAAPPAPAAAPAGPRA-----PPAAVAARE-SP-----214
Qy 293 IMYKGRVTLOKVGHPSCTFLYGPDPFAVRATDPPQVAFPSPAELPDQKQ 342
Db 215 -----LPPPPQ-----PAPPQQQQPPPPPPPPQ 241

RESULT 7

US-10-023-523-47
; Sequence 47, Application US/10023523
; Patent No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-523-47

Query Match 5.5%; Score 150.5; DB 12; Length 550;
Best Local Similarity 23.4%; Pred. No. 0.044;
Matches 82; Conservative 20; Mismatches 111; Indels 137; Gaps 15;
Qy 2 ALAPERAAPRVLFGEWLLGSISSGCVGEGLOWLDEARTCFRVPWKHFARKDLSADARIFK 61
Db 20 AAAASSAASPHYQEWILDTI-----DSLRS-----RKARPD-ERICMVR 60
Qy 62 AWAVARGWPPSSRGSGPPPEAETA-----RAGWKTNFRCAL--RSTRFVMLRDNS 112
Db 61 -----RRHGPERTEAELEKLIQRAVLRSYKGSISYRNAARVQPPRGA 107
Qy 113 GDPADPHKVYALSRELWCWREGPGTDQTEAEAPAAVPPPGGPGPFLLAHTHAGLQAPG 172
Db 108 TPAPPAPR-----RGSP-----AAAAAAPPPTAPPPP-----PAPV 139
Qy 173 PAPAGDKDLLLLQAVQOSCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGGLPA 232

Db 140 AAAA-----APARAPRAAAAAAATAAPPSPGPAQPG 170
Qy 233 GELYGWAVETTPSPGPOPAALTTGEAAAPESPHQAEPLYLSPSPSACTAVQEPSGALDVT 292
Db 171 PR-----AQRAAPLAAPPAPAAAPAGPRA-----PPAAVAARE-SP-----214
Qy 293 IMYKGRVTLOKVGHPSCTFLYGPDPFAVRATDPPQVAFPSPAELPDQKQ 342
Db 215 -----LPPPPQ-----PAPPQQQQPPPPPPPPQ 241

RESULT 8

US-10-162-223-7
; Sequence 7, Application US/10162223
; Publication No. US20030033637A1
; GENERAL INFORMATION:
; APPLICANT: OISHI, KAREN K.
; APPLICANT: COMARATTA, LEONARD M.
; TITLE OF INVENTION: GENE EXPRESSION AND PRODUCTION OF TGF-B PROTEINS
; TITLE OF INVENTION: INCLUDING BIOACTIVE MULLERIAN INHIBITING SUBSTANCE FROM
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 048281/0140
; CURRENT APPLICATION NUMBER: US/10/162,223
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-223-7

Query Match 5.4%; Score 148.5; DB 9; Length 560;
Best Local Similarity 23.7%; Pred. No. 0.06;
Matches 106; Conservative 24; Mismatches 150; Indels 167; Gaps 21;
Qy 64 AVARGWPPSSRGSGPPPEAETAERAGWKTNFRCALRSTRF-VMLRDNSSGDPADPHKV 122
Db 86 AVQARWGP--RDLATFGVCNTGDR-----QAALPSLRLLGAWLR---DFGGQRLVV 132
Qy 123 ALSRELWCWREGPGTDQTEAEAPAAVPPPGG--GP-----PGFLLAHTHAGLQAPG 170
Db 133 LHLEEVTEWPTPSLRPQE-----PPPGAGPPELALLVLPGPGEVTVTRAGL--PG 183
Qy 171 -----PLPAPAGD-KGDLLOAVQOSCLADHLLTASWG-----203
Db 184 AOSLCPSRDRTRYLVLAVDPAAGAWRGSGLALTLPQREGDSRUSTARLQALLFGDDHRCFT 243
Qy 204 -----DPVPTKAPGEGQGLPLTGACAGGGLPAGELYGWAVETTPSPGPQPA 251
Db 244 RMTFALLLPSEPAFLPAHQ-----LDTVFPPRPS 277
Qy 252 ALTTGEAAAPESPHQAEPLYLSPSPSACTAVQEPSGALDVTIMYKGRVTLOKVGHPSCT 311
Db 278 -----AELEESPPSADPFLETLTRLVRALRVP-----PARASAPRLALDPA- 319
Qy 312 FLYGPDPFAVRATDPOOV-----APSPAEPLDQKRLRYTEELLRH 352
Db 320 -LAGFPQGLVNLSDPAALERLLDGEPEPLLLLRRTAATTGDPAPLHDPTSPATWALARR 378
Qy 353 VAPGLH---LELRG-----PQLWARRMGKCKVYWEVGGPPGSPASPTACILPRNCD 401
Db 379 VAAELQAAAELSLPGLPPATAPLLARLLALCP-----GGPGLGDLPLRALLLKA--- 430
Qy 402 TPIDFRVFFQEL-VEFRARQRGSPR 427
Db 431 -----LQGLRVEWRGDRPRGCR 448

RESULT 9

US-09-813-398-23
; Sequence 23, Application US/09813398

```
; Patent No. US20020169292A1
; GENERAL INFORMATION:
; APPLICANT: Bruce D. Weintraub
; APPLICANT: Mariusz W. Szkludinski
; APPLICANT: University of Maryland
; TITLE OF INVENTION: CYSTINE KNOT GROWTH FACTOR MUTANTS
; FILE REFERENCE: UOFWD.003C1
; CURRENT APPLICATION NUMBER: US/09/813,398
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: PCT/US99/05908
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/US98/19772
; PRIOR FILING DATE: 1998-09-22
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 561
; TYPE: PRT
; ORGANISM: HOMO SAPIEN
US-09-813-398-23

Query Match          5.4%; Score 148.5; DB 9; Length 561;
Best Local Similarity 23.6%; Pred. No. 0.06;
Matches 106; Conservative 22; Mismatches 150; Indels 171; Gaps 21;

QY 64 AVARGRWPPSSRGCGPPPEAEAEAEAAVPPPOG-CP-----PCPFLAHTHAGLQARG 170
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 87 AVORARWGP--RDLATFGVCNTGDR-----QAALPSLRRLGAWLR-----DPGGQRLVV 133
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 123 ALSRELCHREGCGTQTEAEAAVPPPOG-CP-----PCPFLAHTHAGLQARG 170
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 134 LHLEEVWTEPTSLRFQE-----PPPGGAGPPELALLVYPGPGPEVTVTRAGL---PG 184
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 171 -----PLPAPAGD-KGDLILLQAVQOSCLADHLLTASWGA----- 203
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 185 AQSLCPSRDTRYLVILAVDRPAGAWGSGLALTLQRGEDSRSLSTARLQALLFGDDHRCFT 244
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 204 -----DPVPTKAPGCGEGLPLTGACAGGGLPAGELYGWAVETTPSPGPOPA 251
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 245 RMTPALLLRSEAPALPAHQ-----LDTVFPFPPRPS 278
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 252 ALTTEGAAPESPHQAEVYLSPSACTAVQEP-----SPGALDVTIMYKGRV 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 279 -----AELESPSPADPEFLETTLRLVRLRVPPARASAPRIALDPDAL----- 321
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 301 LQKVGHPSCTLYGPP-----DPAVRATDQVAFPSPAELPDQKQLRYTEELL 350
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 322 -----AGFPQGLVNLSDPAALERLLDGEELPILLLRPTAATTGDPADLHPTAPWATALA 377
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 351 RHVAFGLH---LELRG-----PQLWARRMGKCKVYWEVGGPPGSPSPSTPACLLPRN 399
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 378 RVAAELQAAAEALSLPGLPPATAPLLARLLALCP-----GGPGGLGDLRALLLKKA- 431
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 400 CDTPIFDVRFVQEL-VEFRARQRRGSPR 427
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 432 -----LQGLRVEWGRDPRGCR 449
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-10-122-805-4
; Sequence 4, Application US/10122805
; Publication No. US20030082645A1
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
; APPLICANT: Porton, Barbara
; APPLICANT: Kao, Hung-Teh
; TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE AND USES
; FILE REFERENCE: 600-1-202 CIP
; CURRENT APPLICATION NUMBER: US/10/122,805
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/129,668
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-05

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/906,865
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-122-805-4

Query Match          5.4%; Score 148; DB 9; Length 696;
Best Local Similarity 27.1%; Pred. No. 0.082;
Matches 73; Conservative 12; Mismatches 94; Indels 90; Gaps 14;

QY 133 GPGTDQTEAEAPA-AVPPPGQGGPPGPFPLAHTHAGLQAPG-----LPAPAGDKGDLILLQAV 187
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 445 GRTSQQPAGPPAQOEPPPPQGGPPQ-----GPGPQRCQPLQQRPPPPQCGQHLSGLGPP 499
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 188 QOSCLADHLLTASWCADVPVTKAP-----GSGQEGPLTGACAGGGLPAGEL 235
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 500 AGSPLPQRL-----ESPTSAPQQPASQAAPPTQGGQGRSPV-----AGGFGAPP----- 544
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 236 YGMAVETTPSPGQ-----PAALTTEAAAPESP-----HQAEPYLSPPSPSACT 279
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 545 ---AARPPASPSQQAQAGPPQATROTSTVSGEAPPKASGAPPQGGQRCQRPKPPGPGAGPT 601
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 280 -----AVQESP-----GALDVTIMYKGRVTLQKVGHPSCTLYGPPD 318
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 602 RQASQAGVPRTGPTTQPPRPSGPGPAGAPKPAQAKP-----SQDVPPTATAAGGPP- 656
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 319 PAVRATDQVAFPSPAELPDQKQLRYTE 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 657 -----HPQLFNLEPEA---PPRPSLSQDE 677
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-10-060-036-159
; Sequence 159, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-159

Query Match          5.4%; Score 148; DB 9; Length 1464;
Best Local Similarity 23.4%; Pred. No. 0.18;
Matches 80; Conservative 11; Mismatches 89; Indels 162; Gaps 17;

QY 66 ARGWRPPSSRGCGPPPEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAE 125
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 414 ARGSPGPGQGGPPGPP-----KGNSEGPAP----- 439
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 126 RELCWRGEGTDTQTEAEAPAAVPPPOG---GPPGPFPLAHTHAGLQA-PGP---LPAPAGDKG 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 440 -----GSKGDTGAKGE-----PGVGVQGGPPGAGEGKRGARGEGPGTGLPFGPGERG 488
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 181 DLLLQAVQOSCLADHLLTASWCADVPVTKAPGSGQEGPLTGACAGGGLPAGELYGWAV 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 489 -----GPCSRGFFGADGVAGPKG-PAGE----- 510
QY 241 ETTSPGCPAALTGTEAAPP-----ESPHQAEPLSPSPSACTAVQEPSGA 288
Db 511 --RGSPGAPGKSPGEGAGPGAGLPGAKGLTGSPG-----SPGPDGKTG--PPGPAG 560
QY 289 LDVTIMYKGRVTLQKVGHSPCTFLYGPDPFAVRATDPQOVAFPPSPAELPDQKQLRYTEE 348
Db 561 QD-----GRP-----GPPGPGARGQAGVMGPPGPKGAAGE----- 591
QY 349 LLRHVAPGLHLELRGPQLWARRMKCKVYWEVGGPPGSPASPS 390
Db 592 -----PGKAGE-RG-----VPGPPGAVGPA 610
RESULT 12
US-10-171-311-36
; Sequence 36, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerseh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-36

Query Match 5.4%; Score 148; DB 9; Length 1464;
Best Local Similarity 23.4%; Pred. No. 0.18;
Matches 80; Conservative 11; Mismatches 89; Indels 162; Gaps 17;
QY 66 ARGWPPSSRGSGPPPEAETAEAGWKTFRCLRRTRRRFVLRDNRSGDPADPHKVALS 125
Db 414 ARGSPGPGGPGPPG-----KGNSEGEFGAP----- 439
QY 126 RELCWREGPGCTDQTEAEAPAAVPPPOG--GPPGPFLLAHTHAGLQA-PGP--LPAPAGDKG 180
Db 440 -----GSKGDTGAKGE-----PGPVGVQPPGAGEEGKRGARGEPGPTGLPGPPGERG 488
QY 181 DLLLQAVQOSCLADHLLTASWGADVPVTKAPGEGQEGPLTGACAGPGPLPAGELYGWAV 240
Db 489 -----GPCSRGFFGADGVAGPKG-PAGE----- 510
QY 241 ETTSPGCPAALTGTEAAPP-----ESPHQAEPLSPSPSACTAVQEPSGA 288
Db 511 --RGSPGAPGKSPGEGAGPGAGLPGAKGLTGSPG-----SPGPDGKTG--PPGPAG 560
QY 289 LDVTIMYKGRVTLQKVGHSPCTFLYGPDPFAVRATDPQOVAFPPSPAELPDQKQLRYTEE 348
Db 561 QD-----GRP-----GPPGPGARGQAGVMGPPGPKGAAGE----- 591
QY 349 LLRHVAPGLHLELRGPQLWARRMKCKVYWEVGGPPGSPASPS 390

Db 592 -----PGKAGE-RG-----VPGPPGAVGPA 610
RESULT 13
US-10-216-705-21
; Sequence 21, Application US/10216705
; Publication No. US20030096973A1
; GENERAL INFORMATION:
; APPLICANT: Meristem Therapeutics, S.A.
; TITLE OF INVENTION: Recombinant Collagens and Derived Proteins Produced by Plants, Mei
; TITLE OF INVENTION: obtaining Such and Their Uses
; FILE REFERENCE: 1149-3 DIV
; CURRENT APPLICATION NUMBER: US/10/216,705
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 09/331,347
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-216-705-21

Query Match 5.4%; Score 148; DB 9; Length 1464;
Best Local Similarity 23.4%; Pred. No. 0.18;
Matches 80; Conservative 11; Mismatches 89; Indels 162; Gaps 17;
QY 66 ARGWPPSSRGSGPPPEAETAEAGWKTFRCLRRTRRRFVLRDNRSGDPADPHKVALS 125
Db 414 ARGSPGPGGPGPPG-----KGNSEGEFGAP----- 439
QY 126 RELCWREGPGCTDQTEAEAPAAVPPPOG--GPPGPFLLAHTHAGLQA-PGP--LPAPAGDKG 180
Db 440 -----GSKGDTGAKGE-----PGPVGVQPPGAGEEGKRGARGEPGPTGLPGPPGERG 488
QY 181 DLLLQAVQOSCLADHLLTASWGADVPVTKAPGEGQEGPLTGACAGPGPLPAGELYGWAV 240
Db 489 -----GPCSRGFFGADGVAGPKG-PAGE----- 510
QY 241 ETTSPGCPAALTGTEAAPP-----ESPHQAEPLSPSPSACTAVQEPSGA 288
Db 511 --RGSPGAPGKSPGEGAGPGAGLPGAKGLTGSPG-----SPGPDGKTG--PPGPAG 560
QY 289 LDVTIMYKGRVTLQKVGHSPCTFLYGPDPFAVRATDPQOVAFPPSPAELPDQKQLRYTEE 348
Db 561 QD-----GRP-----GPPGPGARGQAGVMGPPGPKGAAGE----- 591
QY 349 LLRHVAPGLHLELRGPQLWARRMKCKVYWEVGGPPGSPASPS 390
Db 592 -----PGKAGE-RG-----VPGPPGAVGPA 610

RESULT 14
US-10-149-352-2
; Sequence 2, Application US/10149352
; Publication No. US20030105050A1
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajinder
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; FILE REFERENCE: 06275-254US1
; CURRENT APPLICATION NUMBER: US/10/149,352
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/GB00/04741
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: GB 9929487.8
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 4.0
; SEQ ID NO 2
; LENGTH: 1464
; TYPE: PRT

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:45:13 ; Search time 22 Seconds
(without alignments)
2197.982 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAAPRVLFGEWLLG.....SANSLYDDIECFMLEQPA 503
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	780	28.6	491	2 S56753	interferon regulat
2	482.5	17.7	504	2 G02474	interferon regulat
3	440	16.1	459	2 JC6520	interferon regulat
4	376	13.8	450	2 S57837	lymphoid-specific
5	333.5	12.2	424	2 A35861	interferon consens
6	328.5	12.0	393	2 A45017	transcription fact
7	306	11.2	425	2 A45064	interferon consens
8	259.5	9.5	399	2 JC4592	transcription fact
9	224	8.2	203	2 S57836	lymphoid-specific
10	183	6.7	1334	2 T50568	probable multi-dom
11	176	6.4	325	2 I52998	interferon regulat
12	176	6.4	1146	2 A38587	collagen, cornea-s
13	167	6.1	349	2 A53340	interferon regulat
14	166	6.1	325	2 B31595	interferon regulat
15	166	6.1	329	2 A31595	interferon regulat
16	165.5	6.1	328	2 A36330	interferon regulat
17	160.5	5.9	1400	2 T31555	hypothetical prote
18	160	5.9	1747	1 A45974	collagen alpha 1(X
19	159.5	5.8	416	1 SKX1AG	dermal gland prote
20	159	5.8	576	2 T36729	probable serine/th
21	158.5	5.8	839	2 T04859	extensin homolog F
22	157.5	5.8	705	2 A35363	synapsin I splice
23	156.5	5.7	1464	2 S59856	collagen alpha 1(I
24	156	5.7	381	2 T27806	hypothetical prote
25	156	5.7	1857	2 S31212	collagen alpha 1(X
26	156	5.7	1888	2 S78476	collagen alpha 1(X
27	156	5.7	3149	1 Q0B88	BPLF1 protein - hu
28	155	5.7	801	2 T29018	hypothetical prote
29	154	5.6	660	1 Q0B53	BHLF1 protein - hu

30	154	5.6	3530	2 A59266	unconventional myo
31	150.5	5.5	691	2 A25704	synapsin I - rat
32	150	5.5	704	2 A30411	synapsin Ia - rat
33	149.5	5.5	1952	2 T48814	hypothetical prote
34	148.5	5.4	560	1 WFHUM	mullerian inhibiti
35	148	5.4	384	2 H70580	hypothetical prote
36	148	5.4	1464	1 CGHU1S	collagen alpha 1(I
37	148	5.4	1466	1 CGHU7L	collagen alpha 1(I
38	147	5.4	303	2 T21314	hypothetical prote
39	147	5.4	631	2 B40983	collagen alpha 1(X
40	146.5	5.4	1042	1 CGCH1S	collagen alpha 1(I
41	145	5.3	296	2 A31219	collagen 1 - Caeno
42	145	5.3	1207	2 T00378	KIAA0641 protein -
43	145	5.3	2715	2 T13049	eyelid - fruit fly
44	144.5	5.3	1315	2 A56101	collagen alpha 1(X
45	144.5	5.3	1546	1 CGHU2E	collagen alpha 2(X

ALIGNMENTS

RESULT 1

S56753
interferon regulatory factor 3 - chicken
C:Species: Gallus gallus (chicken)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S56753
R:Grant, C.B.; Vasa, M.Z.; Deeley, R.G.
Nucleic Acids Res. 23, 2137-2146, 1995
A:Title: CIRP-3, a new member of the interferon regulatory factor (IRF) family that is r
A:Reference number: S56753; MUID:95334365; PMD:7541908
A:Accession: S56753
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-491 <GRA>
A:Cross-references: EMBL:U20338; NID:G790580; PIDN:AAA86995.1; PID:G790581
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 28.6%; Score 780; DB 2; Length 491;
Best Local Similarity 38.5%; Pred. No. 3.5e-40;
Matches 208; Conservative 60; Mismatches 180; Indels 92; Gaps 17;

QY	2	ALAPERAAPRVLFGEWLLGEISGCGVEGLWDLEARTCFRVPKWFKAROLDSEADARIFK	61
DB	3	ALDSEGAQKLRFGPWLNAVSSGLYRGLCWIDPPRRIFRIPKWKNARKDVTSDDVEIFK	62
QY	62	ANAVARGRWPPSSRGSGPPPEAETAEACWKTNFCALRSTRFRVLMRNSGDPADPHKV	121
DB	63	AWAKASGRY-----EGNAEDPAKWKTNFCALRSTHMFMLLEDKSCNDDPHKV	111
QY	122	YALSRELWCWREGFGTDTAEAPAAVPPQP--GGP-----PGPFLAHTAGLQ----	167
DB	112	YAVA-----SGVPNDRSGGVPAGALQQQPOLLNHHDLALENTPT	152
QY	168	-APGFLPAPAGDKGL-LLQAVQOCSLADHLITASWADPVPYKAPGEG--QEGFLPTG-	222
DB	153	DSTEGVAAALATQVLDLQLQSVLQHCNISAL-----GSQFTLWAHTGDALPEDALLPGQ	207
QY	223	-ACAGGGLPAGELYCM-AVETTPSPGPOP-----AALTGTGAAAPESPQOAE	269
DB	208	DGCLFGP-----QFQDWRQLEPPLLGNQPLTGGCGGCGQAGALPVSECAIAPSPABE	262
QY	270	YL-----SPSPSACTAVQEPSGALDVTIMYKGRITVLQKVGVHPSCTFLXGPPDPVAVRATD	325
DB	263	LLFQSANPAPPAGDIGGLPLLDITIIYRGWVYQEQVDDSRCLVAYQLDPAV--AE	320
QY	326	POQVAPPSPAELPDQKQLRYTELLRHVAPGLHLEURGPOLWARRMGKCKVYHEV-----	380
DB	321	QRLVLFPSPASLPDPRQRYTEDLLE--VAGLRLEQRAGQLLATRLKCKKCVFVALSQOLE	378
QY	381	GGPPGCSASTPACLLPRNCDTPIDFRVFFQELVEFRARORGRSPRYTIYLGFGGDLSA	440
DB	379	GGEP-----PUNLHRDQETTFIDFRVFCTELDRDSDRRSRRSPDITFLFCGQCQFSS	431

A;Title: An interferon regulatory factor-related gene (xIRF-6) is expressed in the postnatal liver
A;Reference number: JC6520; MUID:98086218; PMID:9426249

A;Contents: Embryo
A;Accession: JC6520

A;Molecule type: mRNA

A;Residues: 1-459 <HAT>

A;Cross-references: DDBJ:D86492; NID:g2780409; PIDN:BAA24349.1; PID:g2780410

C;Comment: This protein plays a role in the transcriptional regulation of specific genes

C;Genetics:

A;Gene: xIRF-6

C;Superfamily: lymphoid-specific interferon regulatory factor

Query Match 16.1%; Score 440; DB 2; Length 459;
Best Local Similarity 27.1%; Pred.No. 1.3e-19;
Matches 140; Conservative 74; Mismatches 197; Indels 106; Gaps 19;

Qy 1 MALAPERAAPVLFGWELLGLBISSCYEGLOWLDEARTCFRVPWKHFAR-KDLSEADARI 59
||| : : : : :
Db 1 MAHPR---RVRLKPLVAOVDGSMYGLIWLNRKAERFQIPWGIAHRSPQEENIT 56
||| : : : : :

Qy 60 FKAVAVAGRWPPSRGGGPPEATBTRAGHKTFRCALSTRFRVLMRDNSGD-PADP 118
||||| : : : : :
Db 57 FKAWAVETGYQEG-----ADEPDPAKWAKLRCAINLKSRFKMLYGDTKEYPMPNP 107
||| : : : : :

Qy 119 HKVTALSRELWCWRPGDGTQTAEAPAAVPPPQQGPGPFALAHTHAQLAQAGLPAPAGD 178
||| : : : : :
Db 108 VKIV----EVC-----DIPOSQG-----SIINPGSTGVSPWD 135
||| : : : : :

Qy 179 KGDLLLLQAQQOSCLADHLLLTASMGADVPTTKAP-----GEOGEGLFLTGACAGGPLPA 232
||| : : : : :
Db 136 DDDFEADELNQS-----ONHVPISEFFNCLINDSPIGSSSTGSCT----- 176
||| : : : : :

Qy 233 GELYGWAVEITPSPCQPQAL--TTGEAAAPESSHQAEPYLSPSPSACTAVQEPSGALD 290
||| : : : : :
Db 177 -----PEQTPKTEPQMEVPTSPPDGADFESP---EMWISSLPWNT-----DLE 217
||| : : : : :

Qy 291 VTIMYKGRTVLQKV-VGHP-SCTFLYGGDPDAVRATD-----POQVAFPSPAELPDQKK 342
||| : : : : :
Db 218 IQFYRGKMGQTMVSNPQCGRFLPYGLGPMNQBELFGPITLEQVRFEGTEQIVNEKQ 277
||| : : : : :

Qy 343 LRYTEEILRHVAPGLHLRGPOLWARMKVKYVWEVGPPGPSASPSTPACLPRNCDF 402
||| : : : : :
Db 278 KLFTSLRLDVNDRGILILEVSGHAIYAIRCQCKVYW---SGCSPSPITPN-FIEQRKV 333
||| : : : : :

Qy 403 PIFDRVFVQBLVEFRARRRRRSPLYTYLGFDQLSAGRPKEXSLVLVKLEPWLKRCHL 462
||| : : : : :
Db 334 KLFCEVFLSDLISHQGIITKPYYEILYGFGEEMPDKYKERKLIIIQIPIVARMII 393
||| : : : : :

Qy 463 EGTORGVSISLDSDLDLCLSANSLYDDIECFMWEL 499
||| : : : : :
Db 394 EMFTGDSRPSDGSIRLIQISIPD-IKDNIYSHLKHLL 429
||| : : : : :

RESULT 4
S57837

Lymphoid-specific interferon regulatory factor - mouse
N;Alternate names: Pip PU.1 interaction partner
C;Species: Mus musculus (house mouse)
C;Date: 28-Oct-1995 #sequence revision 03-Nov-1995 #text_change 29-Sep-1999
C;Accession: S57837; S57838; J49359
R;Matsumura, T.; Grossman, A.; Mitrucecker, H.W.; Siderovski, D.P.; Kleifer, F.; Kawakami
Nucleic Acids Res. 23, 2127-2136, 1995
A;Title: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon regulato
A;Reference number: S57836; MUID:95334364; PMID:7541907
A;Accession: S57837
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-450 <MAT1>
A;Cross-references: EMBL:U20949; NID:g972947; PIDN:AAA75316.1; PID:g972948
A;Note: The nucleotide sequence was submitted to the EMBL Data Library, February 1995
A;Note: Only a part of the coding sequence is given
R;Matsumura, T.; Grossman, A.; Mitrucecker, H.; Siderovski, H.; Siderovski, D.; Kawakami, T.
submitted to the EMBL Data Library, June 1994


```
QY 76 GGGPPPEAETARAGWKTFRFCALRSTRFVMLRDNSSG-DPADPHKYVALSRELWCWREGP 134
Db 70 -----KEGDTGGPAVWKTRLCALNKSSEKVEYPERGRMDVABPYKYQLLPGLVSGQP 124
QY 135 GTQDTEAEAPAAVPPGPGPPFLAHTHAGLQAPGLPAPAGDKGLLQAVQSCSLAD 194
Db 125 GTQKVPSEK-----RQHSVS-----SERKEED-----AMQNCTLSP 156
QY 195 HLLTASGADVPVTKAPGEGQEGPLTGCACAGGGLPAGELYGWAVETTPSGPPAALT 254
Db 157 SVLQDSLNNEE-----EGASG-----GAVHSDIG-----SSSSSSPEPQEVTD 196
QY 255 TGEAAAPESPHQAE-----PYLSPSPSACTAVQSPSG-ALDVTIMYKGRVLOKVVGHPS 309
Db 197 T-----EAFQGDQDSLEFLP-----PEPYSLLFTIYNGRVVGAQVQSOLD 240
QY 310 CTFLYGPDPFAVRATDPOQVAFSPSAPLPPQKQLRYTEELRHVAPGLHLELRGPOLWAR 369
Db 241 CRLV---AEPSSGESSMEQVLPKPGPLEP-----TORLLSQLERGILVASNPRGLFVQ 291
QY 370 RMGCKKYVWVGPPGPGSASTPACLLPRNCDDPIFDPRVFFQELVEFRARQRRG-SPRY 428
Db 292 RLCPPIPSWNAPOAPPQPGPH-----LLPSNECVELFTAYFCRDLV--RYFQGLGPPPKF 345
QY 429 TIVLFGQDLSAGRPKEKSLVVKLEPWLCKRVHLEGT--OREGVSSL 473
Db 346 QVTINFWEEHSGSHSTQNLITVKMEQAFARYLLEQTPEQQAAILSL 392

RESULT 7
A45064
interferon consensus sequence binding protein, ICSBP (DNA-binding domain) - human
C;Species: Homo sapiens (man)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1998
C;Accession: A45064
R;Weisz, A.; Marx, P.; Sharf, R.; Appella, E.; Driggers, P.H.; Ozato, K.; Levi, B.Z.
J. Biol. Chem. 267, 25589-25596, 1992
A;Title: Human interferon consensus sequence binding protein is a negative regulator of
A;Reference number: A45064; MUID:93094284; PMID:1460054
A;Accession: A45064
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-425 <WEI>
A;Experimental source: lung and blood
A;Note: sequence extracted from NCBI backbone (NCBI:P:120312)
C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; transcription regulation

Query Match 11.2%; Score 306; DB 2; Length 425;
Best Local Similarity 23.4%; Pred. No. 1.6e-11;
Matches 121; Conservative 67; Mismatches 174; Indels 156; Gaps 21;

QY 16 EWLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPPSS 74
Db 12 QWLIEQIDSSMYPGLIWENEKSMFRIPWKHAGQDYNQEVDSIFKAWAVFKGK- 67
QY 75 RGGPPPEAETARAGWKTFRFCALRSTRFVMLRDNSSG-DPADPHKYVALSREL- 128
Db 68 -----KEGDKAEPATWKTRLCALNKSDFEEVTDRLSDISPYKYRIVPEEDQCK 121
QY 129 -----CWREGPQDTEAEAPAAVPPGPGPPFLAHTHAGLQAPGLPAPAGDKGD 181
Db 122 LGVATACVNEVTEMECGREIDELIKEPS-----VDDYGMKIKSPSPDA- 168
QY 182 LLLQAVQSCSLADHLLTASGADVPVTKAPGEGQEGPLTGCACAGGGLPAGELYGWAVE 241
Db 169 -----CRSQ--LLPDW-----WAHE 181
QY 242 TTSPGPPQAALTGEAAAPESPHQAEPYLSPSPSACTAVQSPSGALDVTIMYKGRVTL 301
Db 182 --PSTGRR--LVTGYTTY-DAHHSAP-----SQMVISFYGGKLVG 217
```

```
QY 302 QKVGHGHP-SC-----TFLYGPDPFAVRATDPOQVAFSPSAPLPPQKQLRYTEEL 349
Db 218 QATTTCPEGCRSLSLSGPLPGTKLYGP-----EGLELVRFPADTIPSEGRQVTRKL 270
QY 350 LRHVAFG-LHLELRGPOLWARRMKCKVYWEVGGPGSGASPSPTACLLPRNCDDTPIPDFR 408
Db 271 FGLERGLVHSSRQG--VFVKRLCOGRVFCGNAVVCGRPNK-----LERDEVVQVFDTS 324
QY 409 VFFQELVEFRARORRGSPRYTIVLFGQDLSAGRPKEKSLVVKLEPWLCKRVHLEGTQRE 468
Db 325 QFFRELQOQFYNSQGR-LPDGRVVLFCGEEPPDMAPLRSLKILVQIE-----OLYVRQLAEE 379
QY 469 GVSSLDSSDL-----DLCLSSANSLY 489
Db 380 AGKSCGAGSVWQAPBEPBPDPQVFMFPDICASHQRSFF 417

RESULT 8
JC4592
transcription factor ISGF3 gamma chain - mouse
N;Alternate names: interferon-stimulated gene factor 3 gamma chain
C;Species: Mus musculus (house mouse)
C;Date: 10-Apr-1996 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
C;Accession: JC4592; S71599
R;Suhara, W.; Yoneyama, M.; Yonekawa, H.; Fujita, T.
J. Biochem. 119, 231-234, 1996
A;Title: Structure of mouse interferon stimulated gene factor 3 gamma (ISGF3 gamma/p48)
A;Reference number: JC4592; MUID:97037063; PMID:8882710
A;Accession: JC4592
A;Molecule type: mRNA
A;Residues: 1-399 <SUH>
A;Cross-references: EMBL:U51992; NID:g1263309; PIDN:AAC52494.1; PID:g1263310
A;Experimental source: L929 cells
R;Kawakami, T.; Matsumoto, M.; Sato, M.; Harada, H.; Taniguchi, T.; Kitagawa, M.
FEBS Lett. 358, 225-229, 1995
A;Title: Possible involvement of the transcription factor ISGF3gamma in virus-induced exp
A;Reference number: S71599; MUID:95145714; PMID:7843405
A;Accession: S71599
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-87 <KAW>
C;Genetics:
A;Gene: Isgf3g
A;Map position: 14
C;Function:
A;Description: responsible for specific interaction with the promoter element, interferon
ed gene factor 3, the primary regulator of type I interferon responses; involved in type
A;Note: induced by interferon-alpha and interferon-beta
C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; signal transduction; transcription factor
F;10-117/Domain: DNA binding #status predicted <DNB>

Query Match 9.5%; Score 259.5; DB 2; Length 399;
Best Local Similarity 22.5%; Pred. No. 9.6e-09;
Matches 105; Conservative 64; Mismatches 192; Indels 105; Gaps 14;

QY 17 WLIGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPPSSR 75
Db 15 WIVEQVESGHPGVCVDDAAKTMFRIFRPHAGQDFREDQDAALFKAWALFKEK----- 68
QY 76 GGGPPPEAETARAGWKTFRFCALRSTRFVMLRDNSSG-DPADPHKYV-----ALSREL 128
Db 69 -----HXKGDGLGHVAVWKTRLCALNKSSEFEEVPERGRMDVABPYKYRILPAGTLNPQ 124
QY 129 CWREGPQDTEAEAPAAVPPGPGPPFLAHTHAGLQAPGLPAPAGDKGLLQAVQ 188
Db 125 RNQKSPCKRSISCVSPERBENMENGRTNGVNVHSDSGSNITGG-----GGNGS----- 171
QY 189 QSCILADHLLTASGADVPVTKAPGEGQEGPLTGCACAGGGLPAGELYGWAVETTPSPGP 248
Db 172 -----NRSDSNCSNLEEG-----AG----- 189
QY 249 QPAALTTEAAAPESPHQAEPYLSPSPSACTAVQSPSGALDVTIMYKGRVTLQKVVGHP 308
```

```
Db 190 -----TTEATIREDPVFLEHQLPLN-----SDYSLLTFTYGGVVGVKQTVHSL 233
Qy 309 SCTFLYGGPDPAVRATDPOQVAPPSPAEPLDQKQRYTEELRHVAPGLHLRLGQPLWA 368
Db 234 DCLVAERSD---SSSMEOVEFPKP-----DPLEPTQHLLNQLDRGVLVASNSRGLEFV 284
Qy 369 RMGKCKVYVEGGPPGSPASPACLLPRNCDTPIFDVRFVQELVEFRARQRRG-SPR 427
Db 285 QRLCPITPISWNAPEAPPGGPH-----LLPSNKCVELFKTTFYCRDLAQY--PQGGPPPK 338
Qy 428 YTIYLGFGDLSAGRPKESLVLKLEPWLCHVHLEGTORGVSLL 473
Db 339 FOATLHFWEESPSSHSQENLTVQMEQAFARHLLLEKIPERKAAL 384

RESULT 9
S57836
Lymphoid-specific interferon regulator factor - mouse (fragments)
C:Species: Mus musculus (house mouse)
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S57836
R:Matsumura, T.; Grossman, A.; Mittrucker, H.W.; Siderovski, D.P.; Kiefer, P.; Kawakami
Nucleic Acids Res. 23, 2127-2136, 1995
A:Title: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon regula
A:Reference number: S57836; MUID:95334364; PMID:7541907
A:Accession: S57836
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-203 <MAT>
A:Cross-references: EMBL:U1192
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 8.2%; Score 224; DB 2; Length 203;
Best Local Similarity 34.8%; Pred. No. 6.7e-07;
Matches 57; Conservative 22; Mismatches 49; Indels 36; Gaps 6;

Qy 16 EWLGEISGCGYEGQLWLDDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPSS 74
Db 4 QWLQIDSGKYPGLWENEEKSVFPIPKHAGKQDYNEEDAALFKAWLPKGF--- 59
Qy 75 RGGPPPEAETAEARAGKTNFRCALSTRFRFVWLNRNSG-DPADPHKYVALSRELWCWREG 133
Db 60 REGIDKPDPT-----WKLRLCALNKSNDFFELVERSOLDISDPYKYRI----- 105
Qy 134 PGTDQTEAEAPAAVPPPGGPPGPELAHAGL---QAPGGLPA 174
Db 106 -----VPEGQRKNIEKLSHLERGLVLMWAPDGLYA 136

RESULT 10
T50568
Probable multi-domain regulatory protein [imported] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T50568
R:Redenbach, M.; Kieser, H.M.; Denapaita, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A:Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
A:Reference number: Z20556; MUID:97000351; PMID:8843436
A:Accession: T50568
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1334 <RDS>
A:Cross-references: EMBL:AL132220; PIDN:CA861705.1
A:Experimental source: strain A3(2)
C:Genetics:
A>Note: SCC75A.05c

Query Match 6.7%; Score 183; DB 2; Length 1334;
Best Local Similarity 28.1%; Pred. No. 0.0015;
Matches 106; Conservative 25; Mismatches 150; Indels 96; Gaps 22;
```

```
Qy 18 LLGETSSGCGYEGQLWLDDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPSSRG 77
Db 146 VLADLPDRTAEAARW--ETRHFEALRARHTAALDLCQAEHSLPELTALCDGH----- 195
Qy 78 GPPPEAETAEER-----AGWKTNFRCALSTRFRFVWLNRNSG-DPA-----D 117
Db 196 -PLDEPLQALRLRALSRSRTAEALAAVEARR--LLADRLGTDGPELRTLHAEALLSPS 252
Qy 118 PHKYVALSRELWCWREGP-----GTDQTEAEAPAAVPPPGGPPGPPFLAHAGL 166
Db 253 PIPTFGRSRTPGWTSPPGSPASGAGAASGTDVASGAGAAGGPDSPASGSPAVAPSGGG 312
Qy 167 QAPGELPAPAGKGDILLQAVQOOSCLADHLLTASWGADVPYTKAPGEGGGLPTGCACAG 226
Db 313 PAPGWWPAPGTAPGSGSTAPPHDTASAAD---TA---PAPGPTSAPGTA---PAAGTAAP 362
Qy 227 GGLG--PA-GELY--GWA--VETTPSPGPQPAALTTGEA-----AAPESPHQAEPLY 271
Db 363 APGTAGPAGTGYAPGTAPVAGTTTAPGTAPAGTAPAGTAPAGTAPAGTAPAGTAPAGT 422
Qy 272 SPSPSACTAV-----QEPSPGALDVTIMYKGRITVLQKVGHPSCTFLYGGPDPAVRAT 324
Db 423 APAPGSTPAGTVPAPGTAPAPGQPA-----DGR---RPVTGPASCT---GPG---AAT 468
Qy 325 DPQQVA-----FPSPA 335
Db 469 PPEAAAAAASAGSAPSPA 485

RESULT 11
I52998
Interferon regulatory factor 1 - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I52998
R:Cha, Y.; Sims, S.H.; Romine, M.F.; Kaufmann, M.; Deiseroth, A.B.
DNA Cell Biol. 11, 605-611, 1992
A:Title: Human interferon regulatory factor 1: intron/exon organization.
A:Reference number: I52998; MUID:93000481; PMID:1382447
A:Accession: I52998
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <RES>
A:Cross-references: GB:L05072; NID:g184648; PIDN:AAA36043.1; PID:g184649
C:Genetics:
A:Gene: GDB:IRF1
A:Cross-references: GDB:127269; OMIM:147575
A:Map position: Sq31.1-5q31.1
A:Introns: 29/3; 63/1; 122/1; 138/3; 182/1; 223/1; 239/3; 285/1

Query Match 6.4%; Score 176; DB 2; Length 325;
Best Local Similarity 23.6%; Pred. No. 0.0089;
Matches 69; Conservative 35; Mismatches 108; Indels 80; Gaps 12;

Qy 11 RVLFGELGELTSSGCGYEGQLWLDDEARTCFRVPWKHFARK--DLSEADARIFKAWAVARG 68
Db 5 RMRMRPWLEMQINSQIPLGLIWINKEEMIFQIPWKHAAKHGWDINK-DACLFERSWAIHTG 63
Qy 69 RNPSPSRGGPPPEAETAEARAGKTNFRCALSTRFRFVWLNRD---NSGDPADPHKYVALS 125
Db 64 RY-----KAGEKEPDPT-----WKNFRCCAMNSLPDIEVKQOQRNKGSSA--VRVYRML 112
Qy 126 RELC---WREGPGTDTQTEAEAPAAVPPPGGPPGPP-----LAHTHAGLQAPGGLPA 174
Db 113 PLTKNQKERSKSRDAKSKAKSKSCGDSDFDPSDGLSSSTLPDDHSSYTVPG--- 168
Qy 175 PAGDKGDLQLQAVQOOSCLADHLLTASWGADVPYTKAPGEGQGLPLTGACAGGGLP--- 231
Db 169 -----YMQDLEVEQALT-----PALSPCAVSSTLPDWH 196
Qy 232 -----AGELYGWAVETTPSPGPQPAALTTGEAAAPESP HQAEPLYSPS 274
Db 197 IPVEVVPDSTSLYNFQV-----SPMPSTSEATTDEDEGKLPEIDIMKLEQS 244
```



```

Db      66 KAGEKEPDPT-----WKANFRCAAMSLPDIIEVKDQSRNKGSSA--VRVYRMLPPLTKN 118
QY      130 -WREGFGTDQTEAEAPAAVPPPGGPPGF-----LAHTHAGLQAPGPLPAPAGDKG 180
Db      119 QRKERKSSRDASKAKRKSCKDSSPDFTSDGLSSSTLPDDHSSYTVFG-----168
QY      181 DILLQAVQOQSLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLP-----231
Db      169 -----YMODLEVEQALT-----PALSPCAVSVSTLPDWHIPVEVV 202
QY      232 ---AGELYGMVAVETTPSPGPQPAALTGFAAAPESPQAEPLSPS 274
Db      203 PDSTSLYNFQV-----SPMPSISEATTDEEGKLPEDIMKLEQS 244

RESULT 15
A31595
interferon regulatory factor 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 05-Nov-1999
C:Accession: A31595
R:Miyamoto, M.; Fujita, T.; Kimura, Y.; Maruyama, M.; Harada, H.; Sudo, Y.; Miyata, T.;
Cell 54, 903-913, 1988
A:Title: Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifica
A:Reference number: A90903; MUID:88311092; PMID:3409321
A:Accession: A31595
A:Molecule type: mRNA
A:Residues: 1-329 <MIY>
A:Cross-references: GB:M21065; NID:gl98458; PIDN:AAA39334.1; PID:9293677; GB:J03160; GB:
C:Keywords: DNA binding; transcription regulation
```

```

Query Match      6.1%; Score 166; DB 2; Length 329;
Best Local Similarity 21.6%; Pred. No. 0.0036;
Matches 65; Conservative 40; Mismatches 88; Indels 108; Gaps 12;

QY      11 RVLFGEWLLGEISSGCGYEGLOWLDEARTCFVPWKHFARK--DLSEADARIFKAWAVARG 68
Db      5  RMRMRPWLEMQINSQIPLGIWINKEMIFQIPWKHAARKHWDINK-DACLFRSWAIHTG 63
QY      69 RWPPSSRGGPPPEAETAEAGWKTNFRCALRSTRFVMLRD--NSGDPADPHKVYAL- 124
Db      64 RY-----KAGEKEPDPT-----WKANFRCAAMSLPDIIEVKDQSRNKGSSA--VRVYRML 112
QY      125 -----SRELWREGFGT-----136
Db      113 PPLTRNQRKERKSSRDTKSKTKKLCGDVSPDFTSDGLSSSTLPDDHSSYTTQGYLQ 172
QY      137 -----DQTEAEAPAAVPPPGGPPGFFLAHTHAGLQ-----APGPLPAPA 176
Db      173 DLDMERDITPALLSPCVV-----SSSLSEWHMQMDIIPDSTTDLYNLQVSPMPSTSEA 224
QY      177 GDKGDLQLQAVQOQSLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELY 236
Db      225 ATDED-----BEGKTAEDIMKLFQESWQPTHIDGKY-----LLNEFGTQLSSVY 270
QY      237 G 237
Db      271 G 271
```

Search completed: June 18, 2003, 12:48:17
Job time : 24 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 16:29:07 ; Search time 163 Seconds
(without alignments)
1193.500 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731

Sequence: 1 MALAPERAAPRVFGWLLG.....SANSLYDDIECFMLEPQA 503

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 385760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2731	100.0	503	2 AAY15103	Aay15103 Modified
2	2719	99.6	503	8 ADJ75462	Adj75462 Marker ge
3	2719	99.6	503	8 ADJ75410	Adj75410 Marker ge
4	2715	99.4	503	7 ADE58665	Ades58665 Human Pro
5	2715	99.4	503	7 ADE58669	Ades58669 Human Pro
6	2695	98.7	503	4 AAE09329	Aae09329 Human int
7	2682	98.2	516	8 ADJ75465	Adj75465 Marker ge
8	2682	98.2	516	8 ADJ75413	Adj75413 Marker ge
9	2682	98.2	516	8 ADL83070	Adl83070 Human PRO
10	2682	98.2	516	8 ADN04327	Adn04327 Antipsori
11	2682	98.2	516	8 ADP24098	Adp24098 PRO Polyp
12	2533.5	92.8	474	8 ADJ75411	Adj75411 Marker ge
13	2533.5	92.8	474	8 ADJ75463	Adj75463 Marker ge
14	2225.5	81.5	441	8 ABM83030	Abm83030 Human dia
15	2210	80.9	440	8 ABM83031	Abm83031 Human dia
16	2177.5	79.7	456	8 ABM83029	Abm83029 Human dia
17	2080	76.2	399	8 ABM83032	Abm83032 Human dia
18	2070.5	75.8	385	7 ADG77002	Adg77002 Human nuc
19	1758.5	64.4	365	8 ABM83033	Abm83033 Human dia
20	1712.5	62.7	342	8 ABM83034	Abm83034 Human dia
21	1567.5	57.4	542	2 AAY15104	Aay15104 Chimeric
22	1562	57.2	457	7 ADE58663	Ades58663 Rat Prote
23	1562	57.2	457	7 ADE58667	Ades58667 Rat Prote
24	1562	57.2	457	8 ADJ76202	Adj76202 Marker ge
25	1562	57.2	457	8 ADJ76237	Adj76237 Marker ge

26	1264.5	46.3	247	8	ADR99903	Adr99903 Immune Re
27	843	30.9	186	8	ADR99902	Adr99902 Immune Re
28	832	30.5	164	8	ADJ75464	Adj75464 Marker ge
29	832	30.5	164	8	ADJ75412	Adj75412 Marker ge
30	832	30.5	164	8	ADR14355	Adri14355 Human NF-
31	497.5	18.2	496	8	ABM83016	Abm83016 Human dia
32	487.5	17.9	476	8	ABM83017	Abm83017 Human dia
33	487.5	17.9	476	8	ABM83019	Abm83019 Human dia
34	487.5	17.9	476	8	ABM83020	Abm83020 Human dia
35	484.5	17.7	488	7	ADD01173	Add01173 Human nuc
36	484	17.7	491	8	ABM83015	Abm83015 Human dia
37	484	17.7	491	8	ABM83018	Abm83018 Human dia
38	469	17.2	455	7	ADG77016	Adg77016 Human nuc
39	469	17.2	455	7	ABM84398	Abm84398 Human dia
40	458	16.8	427	4	AAW78986	Aaw78986 Human pro
41	458	16.8	427	8	ABM82144	Abm82144 Tumour-as
42	458	16.8	427	8	ADQ39628	Adq39628 Human myo
43	458	16.8	427	8	ADQ39630	Adq39630 Human myo
44	454	16.6	427	3	AAW87783	Aay87783 Human IRF
45	454	16.6	427	6	ABG76085	Abg76085 Human int

ALIGNMENTS

RESULT 1

AAV15103

ID AAY15103 standard; protein; 503 AA.

XX AAY15103;

DT 25-JAN-2000 (first entry)

XX Modified Interferon Regulatory Factor-7 protein (IRF-7).

XX Interferon regulatory factor; IRF-7; transactivation domain; serine; threonine; carboxy terminus; aspartic acid; mutant; phosphorylation;

XX post-translational modification; sendai virus; cancer treatment; herpes; activator; promoter; ISRE regulatory element; stimulation; influenza;

XX DNA binding; transcriptional activity; viral infection; HIV infection;

XX activate; homology; cytokine gene; target cell.

XX Homo sapiens.

OS Synthetic.

XX Key

XX Location/Qualifiers

FT Misc-difference 477

FT Misc-difference 479 /note= "Wild type Ser replaced with Asp"

FT Misc-difference 479 /note= "Wild type Ser replaced with Asp"

XX WO9951737-A1.

XX 14-OCT-1999.

XX 07-APR-1999; 99WO-CA000314.

XX 07-APR-1998; 98CA-02234588.

XX (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.

XX Hiscott J, Lin R;

XX WPI; 1999-620201/53.

XX N-PSDB; AAX90993.

XX DR

XX Carboxy-terminus modified highly active forms of interferon regulatory

XX Factor proteins used for the treatment of viral infections.

XX Claim 14; Fig 12; 93pp; English.

XX The present protein sequence is the interferon regulatory factor-7 (IRF-

CC 7) 2D protein, that is modified in the transactivation domain. The serine

CC and threonine residues in the carboxy terminus are modified post-
CC translationally, by phosphorylation, following sendai virus infection.
CC The modified IRF-7, substituted with aspartic acid, functions as a
CC activator of promoters containing ISRE regulatory elements and
CC stimulation of DNA binding and transcriptional activity. IRF-7 protein
CC shows highest homology to IRF-3. The modified IRF proteins are used for
CC the treatment of viral infections like, influenza, herpes or HIV
CC infection. They may also be used to activate a cytokine gene, in cancer
CC treatment or to modify a target cell of an organism
XX
XX
SQ Sequence 503 AA;
Query Match 100.0%; Score 2731; DB 2; Length 503;
Best Local Similarity 100.0%; Pred. No. 1-2e-196;
Matches 503; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALAPERAAAPRVLFGEWLLGEISGGCYEGQLWLDARTCFRVPWKHPARKDLSEADARIF 60
DB 1 MALAPERAAAPRVLFGEWLLGEISGGCYEGQLWLDARTCFRVPWKHPARKDLSEADARIF 60
QY 61 KAWAVARGRWPSSRGSGPPPEAETAEAGKTNFRCALRSTRFRVWMLRNSGDPADPHK 120
DB 61 KAWAVARGRWPSSRGSGPPPEAETAEAGKTNFRCALRSTRFRVWMLRNSGDPADPHK 120
QY 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLPAPAGDKG 180
DB 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLPAPAGDKG 180
QY 181 DILLQAVQOQSLADHLTLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
DB 181 DILLQAVQOQSLADHLTLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
QY 241 ETTSPGQPOAALTGTEAAAPESPHEQAEPLYSPPSACTAVQSPGALDVTIMYKGRTV 300
DB 241 ETTSPGQPOAALTGTEAAAPESPHEQAEPLYSPPSACTAVQSPGALDVTIMYKGRTV 300
QY 301 LQKVVGHPSTFLYGGPDPDAVRATDPOQVAFPPPAELPDQKQRYTEELLRHVAPGLHLE 360
DB 301 LQKVVGHPSTFLYGGPDPDAVRATDPOQVAFPPPAELPDQKQRYTEELLRHVAPGLHLE 360
QY 361 LRGPQLWARRMGKCKVYVEVGPPGSGASPSSTPACLLPRNCDTPIFDPRVFFQELVEPRAR 420
DB 361 LRGPQLWARRMGKCKVYVEVGPPGSGASPSSTPACLLPRNCDTPIFDPRVFFQELVEPRAR 420
QY 421 QRRGSPRYTILYLGFGQDLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSLDL 480
DB 421 QRRGSPRYTILYLGFGQDLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSLDL 480
QY 481 CLSSANSLYDDIECFMLELQPA 503
DB 481 CLSSANSLYDDIECFMLELQPA 503
RESULT 2
ADJ75462
ID ADJ75462 standard; protein; 503 AA.
XX
XX AC ADJ75462;
XX
XX 20-MAY-2004 (first entry)
XX
XX Marker gene related amino acid sequence SEQ ID NO:714.
XX
XX bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.
XX
XX Homo sapiens.
XX
XX EPI394274-A2.
XX
XX 03-MAR-2004.
XX
XX

PF 04-AUG-2003; 2003EP-00254857.
XX
XX 06-AUG-2002; 2002JP-00229312.
PR 20-MAR-2003; 2003JP-00077212.
XX
XX (GENO-) GENOX RES INC.
XX
XX Ontani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX WPI; 2004-193155/19.
DR
XX
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
XX
XX Example 11; SEQ ID NO 714; 241pp; English.
XX
XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC invention is used in the exemplification of the present invention.
XX
XX Sequence 503 AA;
Query Match 99.6%; Score 2719; DB 8; Length 503;
Best Local Similarity 99.6%; Pred. No. 9.5e-196;
Matches 501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALAPERAAAPRVLFGEWLLGEISGGCYEGQLWLDARTCFRVPWKHPARKDLSEADARIF 60
DB 1 MALAPERAAAPRVLFGEWLLGEISGGCYEGQLWLDARTCFRVPWKHPARKDLSEADARIF 60
QY 61 KAWAVARGRWPSSRGSGPPPEAETAEAGKTNFRCALRSTRFRVWMLRNSGDPADPHK 120
DB 61 KAWAVARGRWPSSRGSGPPPEAETAEAGKTNFRCALRSTRFRVWMLRNSGDPADPHK 120
QY 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLPAPAGDKG 180
DB 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTAGLQAPGLPAPAGDKG 180
QY 181 DILLQAVQOQSLADHLTLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
DB 181 DILLQAVQOQSLADHLTLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
QY 241 ETTSPGQPOAALTGTEAAAPESPHEQAEPLYSPPSACTAVQSPGALDVTIMYKGRTV 300
DB 241 ETTSPGQPOAALTGTEAAAPESPHEQAEPLYSPPSACTAVQSPGALDVTIMYKGRTV 300

QY 301 LOKVVGHSCTFLYGPDPVAVRATDQVAFSPAEPLDQKQRYTEELLRHVAPGLHLE 360
DB 301 LOKVVGHSCTFLYGPDPVAVRATDQVAFSPAEPLDQKQRYTEELLRHVAPGLHLE 360
QY 361 LRGPOLWARRMGKCKVYVEVGPPGSGASPSPTACLLPRNCDTPIFDRVFFOELVEFRAR 420
DB 361 LRGPOLWARRMGKCKVYVEVGPPGSGASPSPTACLLPRNCDTPIFDRVFFOELVEFRAR 420
QY 421 QRRGSPRYTIYLGFGQDLSAGRPKEKSLVLKLEPWLKRVHLEGTQREGVSSLDSSLDL 480
DB 421 QRRGSPRYTIYLGFGQDLSAGRPKEKSLVLKLEPWLKRVHLEGTQREGVSSLDSSLDL 480
QY 481 CLSSANSLYDDIECFMELQPA 503
DB 481 CLSSANSLYDDIECFMELQPA 503
RESULT 3
ADJ75410
ID ADJ75410 standard; protein; 503 AA.
AC ADJ75410;
XX
XX
XX 20-MAY-2004 (first entry)
XX
DE Marker gene related amino acid sequence SEQ ID NO:662.
XX
XX bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.
XX
OS Homo sapiens.
XX
XX EP1394274-A2.
PN
XX
PD 03-MAR-2004.
XX
PF 04-AUG-2003; 2003EP-00254857.
XX
PR 06-AUG-2002; 2002JP-00229312.
PR 20-MAR-2003; 2003JP-00077212.
XX
XX (GENO-) GENOX RES INC.
XX
XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX WPI; 2004-193155/19.
XX
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
XX
XX Example 11; SEQ ID NO 662; 241pp; English.
XX
XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic

CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 503 AA;
Query Match 99.6%; Score 2719; DB 8; Length 503;
Best Local Similarity 99.6%; Pred. No. 9.5e-196;
Matches 501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALAPERAAPRVLFCEWLLGHSICCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAPRVLFCEWLLGHSICCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTNFRCALRSTRFVMLRDNSSGDPADPHK 120
DB 61 KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTNFRCALRSTRFVMLRDNSSGDPADPHK 120
QY 121 VVALSRELWCWREGPCTDTEAEAPAAVPPQGGPPGPFPLAHTHAGLQAPGPIAPAGDKG 180
DB 121 VVALSRELWCWREGPCTDTEAEAPAAVPPQGGPPGPFPLAHTHAGLQAPGPIAPAGDKG 180
QY 181 DILLQAVQOQSLADHLLTASWGADVPVTKAPGEGQEGPLTGCACAGGPGLPAGELYGWAV 240
DB 181 DILLQAVQOQSLADHLLTASWGADVPVTKAPGEGQEGPLTGCACAGGPGLPAGELYGWAV 240
QY 241 ETTSPGPGPQPAALTGEEAAAPESPHOABPYLSPPSACTAVQEPSGALDVTIMYKGRV 300
DB 241 ETTSPGPGPQPAALTGEEAAAPESPHOABPYLSPPSACTAVQEPSGALDVTIMYKGRV 300
QY 301 LOKVVGHSCTFLYGPDPVAVRATDQVAFSPAEPLDQKQRYTEELLRHVAPGLHLE 360
DB 301 LOKVVGHSCTFLYGPDPVAVRATDQVAFSPAEPLDQKQRYTEELLRHVAPGLHLE 360
QY 361 LRGPOLWARRMGKCKVYVEVGPPGSGASPSPTACLLPRNCDTPIFDRVFFOELVEFRAR 420
DB 361 LRGPOLWARRMGKCKVYVEVGPPGSGASPSPTACLLPRNCDTPIFDRVFFOELVEFRAR 420
QY 421 QRRGSPRYTIYLGFGQDLSAGRPKEKSLVLKLEPWLKRVHLEGTQREGVSSLDSSLDL 480
DB 421 QRRGSPRYTIYLGFGQDLSAGRPKEKSLVLKLEPWLKRVHLEGTQREGVSSLDSSLDL 480
QY 481 CLSSANSLYDDIECFMELQPA 503
DB 481 CLSSANSLYDDIECFMELQPA 503
RESULT 4
ADE58665
ID ADE58665 standard; protein; 503 AA.
XX
AC ADE58665;
XX
DT 29-JAN-2004 (first entry)
XX
XX Human Protein Q92985, SEQ ID NO 4541.
DE
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
XX WO2003016475-A2.

XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q92985.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a human protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 503 AA;
Query Match 99.4%; Score 2715; DB 7; Length 503;
Best Local Similarity 99.4%; Pred. No. 1.9e-195;
Matches 500; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALAPERAAFRVLFGELLGELISSGCGVEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAFRVLFGELLGELISSGCGVEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAMAVARGRPPSSRRGGPPPEAFTAEARAGKTNFRCAKSTRFRFVRLRNSGDDPADPHK 120
DB 61 KAMAVARGRPPSSRRGGPPPEAFTAEARAGKTNFRCAKSTRFRFVRLRNSGDDPADPHK 120
QY 121 VYALSRLCWRREGGTTQTAEAPAAVPPQGGPPGPFLLATHAGLQAPGFLPAPADKG 180
DB 121 VYALSRLCWRREGGTTQTAEAPAAVPPQGGPPGPFLLATHAGLQAPGFLPAPADKG 180
QY 181 DLLLQAVQSCSLADHLLTASGADPVTKAPGEGQGLPLTGACAGGPGLPAGELYGNAV 240
DB 181 DLLLQAVQSCSLADHLLTASGADPVTKAPGEGQGLPLTGACAGGPGLPAGELYGNAV 240
QY 241 ETTTSPGPQPAALTGEEAAAPESP HQAEPYLSPPSACTAVQSPSPGALDVTIMYKGRTV 300

DB 241 ETTTSPGPQPAALTGEEAAAPESP HQAEPYLSPPSACTAVQSPSPGALDVTIMYKGRTV 300
QY 301 LQKVGHPSCTFLYGPDPFAVRATDPQOAFPSPAELPQOKQLRYTEELRHVAPGLHLE 360
DB 301 LQKVGHPSCTFLYGPDPFAVRATDPQOAFPSPAELPQOKQLRYTEELRHVAPGLHLE 360
QY 361 LRGPQLWARRMGKCKVYWEVGGPPGSPSTPACLLPRNCDTPIFDPRVFQELVEFRAR 420
DB 361 LRGPQLWARRMGKCKVYWEVGGPPGSPSTPACLLPRNCDTPIFDPRVFQELVEFRAR 420
QY 421 QRRGSPRYTIYLGFGQDLGAGRPKEKSLVLKLEPWLRCVHLEGTQREGVSSLDSSDL 480
DB 421 QRRGSPRYTIYLGFGQDLGAGRPKEKSLVLKLEPWLRCVHLEGTQREGVSSLDSSDL 480
QY 481 CUSSANSLYDDIECFLMELEQPA 503
DB 481 CUSSANSLYDDIECFLMELEQPA 503
RESULT 5
AD58669
ID ADE58669 standard; protein; 503 AA.
XX AC ADE58669;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein Q92985, SEQ ID NO 4545.
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q92985.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 503 AA;

Query Match 99.4%; Score 2715; DB 7; Length 503;
Best Local Similarity 99.4%; Pred. No. 1.9e-195;
Matches 500; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALAPERAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWPSSRRGGPPPEAETAEAGWKTNFRCALRSTRFRVWLDRDNGDPADPHK 120
DB 61 KAWAVARGWPSSRRGGPPPEAETAEAGWKTNFRCALRSTRFRVWLDRDNGDPADPHK 120
QY 121 VVALSRELWCWREGPGTDQTEAEAPAAVPPQGGPPGFLAHTHAGLQAPGLPAPAGDKG 180
DB 121 VVALSRELWCWREGPGTDQTEAEAPAAVPPQGGPPGFLAHTHAGLQAPGLPAPAGDKG 180
QY 181 DLLLQAVQOSCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGMV 240
DB 181 DLLLQAVQOSCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGMV 240
QY 241 ETPSPGPOPAALTTCGEAAAPESPHQAEPLYSPSPSACTAVQSPSGALDVTIMYKRTV 300
DB 241 ETPSPGPOPAALTTCGEAAAPESPHQAEPLYSPSPSACTAVQSPSGALDVTIMYKRTV 300
QY 301 LQKVGHPSCTFLYGPDPVAVRATDQVAFPSPAELPDQKQRYTEELLRHVAPGLHLE 360
DB 301 LQKVGHPSCTFLYGPDPVAVRATDQVAFPSPAELPDQKQRYTEELLRHVAPGLHLE 360
QY 361 LRGPQLWARRMGKCKVYWEVGGPPGSGASPSPTACLLPRNCDTPIFDFRVFFQELVEFRAR 420
DB 361 LRGPQLWARRMGKCKVYWEVGGPPGSGASPSPTACLLPRNCDTPIFDFRVFFQELVEFRAR 420
QY 421 QRRGSPRYTIYLGFGQDLSAGRPKESLVLVKLEPWLRCRVHLEGTQREGVSSLDSSDLDL 480
DB 421 QRRGSPRYTIYLGFGQDLSAGRPKESLVLVKLEPWLRCRVHLEGTQREGVSSLDSSSLSL 480
QY 481 CLSSANSLYDDIECFLEMELEQPA 503
DB 481 CLSSANSLYDDIECFLEMELEQPA 503

RESULT 6
AAE09329 standard; protein; 503 AA.
XX
XX AAE09329;
XX
XX
DT 19-NOV-2001 (first entry)
XX Human intracellular regulatory molecule, KWC02.
DE Human; intracellular regulator; cell division; proliferation; therapy;
XX cancer; infection; wound; developmental abnormality; metabolic problem;
KW cytostatic; antibacterial; vulnary; transcription factor; KWC02;
KW interferon response factor; IRF.
XX
OS Homo sapiens.

XX US6274312-B1.
PN 14-AUG-2001.
XX
PD 10-DEC-1997; 97US-00999774.
XX
PF 11-DEC-1996; 96US-0032818P.
XX
PR (SCHE) SCHERING CORP.
XX
PA Gish KC, Seghezzi W, Shanahan F, Lees EM, Mcclanahan TK;
XX WPI; 2001-535086/59.
XX DR N-PSDB; AAD16105.
XX
XX New genes encoding intracellular regulatory molecules, useful for
PT regulating cell division and proliferation (e.g. tumor cells),
PT particularly for treating cancer, infections, wounds, or developmental or
PT metabolic abnormalities.
XX
PS Example 2; Col 33-38; 62pp; English.
XX
XX The present sequence is a human intracellular regulatory molecule, KWC02.
CC The KWC02, an interferon response factor (IRF) homologue, is a
CC transcription factor. The polynucleotides encoding intracellular
CC regulatory molecules are useful for regulating cell division and
CC proliferation of various cell types, including tumour cells.
CC Specifically, they are also useful for treating cancer, infections,
CC wounds, developmental abnormalities or metabolic problems
XX
SQ Sequence 503 AA;
Query Match 98.7%; Score 2695; DB 4; Length 503;
Best Local Similarity 99.0%; Pred. No. 6.1e-194;
Matches 498; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 MALAPERAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAPRVLFGEWLLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWPSSRRGGPPPEAETAEAGWKTNFRCALRSTRFRVWLDRDNGDPADPHK 120
DB 61 KAWAVARGWPSSRRGGPPPEAETAEAGWKTNFRCALRSTRFRVWLDRDNGDPADPHK 120
QY 121 VVALSRELWCWREGPGTDQTEAEAPAAVPPQGGPPGFLAHTHAGLQAPGLPAPAGDKG 180
DB 121 VVALSRELWCWREGPGTDQTEAEAPAAVPPQGGPPGFLAHTHAGLQAPGLPAPAGDKG 180
QY 181 DLLLQAVQOSCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGMV 240
DB 181 DLLLQAVQOSCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGMV 240
QY 241 ETPSPGPOPAALTTCGEAAAPESPHQAEPLYSPSPSACTAVQSPSGALDVTIMYKRTV 300
DB 241 ETPSPGPOPAALTTCGEAAAPESPHQAEPLYSPSPSACTAVQSPSGALDVTIMYKRTV 300
QY 301 LQKVGHPSCTFLYGPDPVAVRATDQVAFPSPAELPDQKQRYTEELLRHVAPGLHLE 360
DB 301 LQKVGHPSCTFLYGPDPVAVRATDQVAFPSPAELPDQKQRYTEELLRHVAPGLHLE 360
QY 361 LRGPQLWARRMGKCKVYWEVGGPPGSGASPSPTACLLPRNCDTPIFDFRVFFQELVEFRAR 420
DB 361 LRGPQLWARRMGKCKVYWEVGGPPGSGASPSPTACLLPRNCDTPIFDFRVFFQELVEFRAR 420
QY 421 QRRGSPRYTIYLGFGQDLSAGRPKESLVLVKLEPWLRCRVHLEGTQREGVSSLDSSDLDL 480
DB 421 QRRGSPRYTIYLGFGQDLSAGRPKESLVLVKLEPWLRCRVHLEGTQREGVSSLDSSSLSL 480
QY 481 CLSSANSLYDDIECFLEMELEQPA 503
DB 481 CLSSANSLYDDIECFLEMELEQPA 503

RESULT 7	Matches	493;	Conservative	2;	Mismatches	2;	Indels	0;	Gaps	0;
ADJ75465	QY	7	RAAPRVLFGEWLLGEISSCYEGLQWLDDEARTCFRVPWKHFARKDI;SEADARIFKAWAVA	66						
ID ADJ75465 standard; protein; 516 AA.	DB	20	RAAPRVLFGEWLLGEISSCYEGLQWLDDEARTCFRVPWKHFARKDI;SEADARIFKAWAVA	79						
AC ADJ75465;	QY	67	RGRWPPSSRGSGPPPEAETAEARAGWKTNFRFCALRSTRFVMLRDNSGDPADPHKVVYALS	126						
DT 20-MAY-2004 (first entry)	DB	80	RGRWPPSSRGSGPPPEAETAEARAGWKTNFRFCALRSTRFVMLRDNSGDPADPHKVVYALS	139						
DE Marker gene related amino acid sequence SEQ ID NO:717.	QY	127	ELCWREGPGTDQTEABAPAAVPPQGGPPGFLLAHTHAGLQAPGPIAPAGDKDILLQA	186						
KW bronchial asthma; chronic obstructive pulmonary disease;	DB	140	ELCWREGPGTDQTEABAPAAVPPQGGPPGFLLAHTHAGLQAPGPIAPAGDKDILLQA	199						
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;	QY	187	VQOSCLADHLLTASWGADPVPTKAPGEGQEGPLTGCACAGGPGLPAGELYGNVAVETPSP	246						
KW gene therapy; marker.	DB	200	VQOSCLADHLLTASWGADPVPTKAPGEGQEGPLTGCACAGGPGLPAGELYGNVAVETPSP	259						
OS Homo sapiens.	QY	247	GPOPAALTTGEEAAAPESPHQAEPLYLSPSPSACTAVQEPSPGALDVTIMYKGRVTLQKVVG	306						
XX EP1394274-A2.	DB	260	GPOPAALTTGEEAAAPESPHQAEPLYLSPSPSACTAVQEPSPGALDVTIMYKGRVTLQKVVG	319						
XX 03-MAR-2004.	QY	307	HPSCFTLYGPPDPVAVRATDPPQVAFPSPAELPDQKQRLRYTEELLRHVAPGLHLELRGPQL	366						
XX 04-AUG-2003; 2003EP-00254857.	DB	320	HPSCFTLYGPPDPVAVRATDPPQVAFPSPAELPDQKQRLRYTEELLRHVAPGLHLELRGPQL	379						
XX 06-AUG-2002; 2002JP-00229312.	QY	367	WABRMGCKKYWEVGGPGSGASPSPTACLLPRNCDTPIEDFRVFFOELVFRARORRQSP	426						
XX 20-MAR-2003; 2003JP-00077212.	DB	380	WABRMGCKKYWEVGGPGSGASPSPTACLLPRNCDTPIEDFRVFFOELVFRARORRQSP	439						
XX (GENO-) GENOX RES INC.	QY	427	RYTYILGFGQDLSAGRPKEKSLVLVLEPWLCHVLEGTQREGVSSLDSDLDLCLSSAN	486						
XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;	DB	440	RYTYILGFGQDLSAGRPKEKSLVLVLEPWLCHVLEGTQREGVSSLDSDLDLCLSSAN	499						
XX WPI; 2004-193155/19.	QY	487	SLYDDIECFLEMELEQPA 503							
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by	DB	500	SLYDDIECFLEMELEQPA 516							
XX comparing the expression level of a marker gene in a biological sample										
XX from a subject with the expression level of the gene in a sample from a										
XX healthy subject.										
XX Example 11; SEQ ID NO 717; 241pp; English.										
XX The present invention describes a method of testing for bronchial asthma										
XX or chronic obstructive pulmonary disease. The method comprises										
XX determining the expression level of a marker gene in a biological sample										
XX from a subject, comparing the expression level determined with the										
XX expression level of the marker gene in a biological sample from a healthy										
XX subject, and judging whether the subject has bronchial asthma or chronic										
XX obstructive pulmonary disease. The marker gene comprises: (a) a group of										
XX genes (S1) whose expression levels increase when respiratory epithelial										
XX cells are stimulated with interleukin-13; or (b) a group of genes (S2)										
XX whose expression levels decrease when respiratory epithelial cells are										
XX stimulated with interleukin-13. Also described: (1) a reagent (I) for										
XX testing for bronchial asthma or chronic obstructive pulmonary disease;										
XX (2) a kit for screening for a candidate compound for a therapeutic agent										
XX to treat bronchial asthma or chronic obstructive pulmonary disease; (3)										
XX an animal model for bronchial asthma or chronic obstructive pulmonary										
XX disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a										
XX method for producing an animal model for bronchial asthma or chronic										
XX obstructive pulmonary disease; (6) a therapeutic agent for bronchial										
XX asthma or chronic obstructive pulmonary disease, comprising the compound,										
XX a marker gene or an antisense nucleic acid corresponding to a portion of										
XX the marker gene, a ribozyme, a polynucleotide that suppresses the										
XX expression of the gene through an RNAi effect or an antibody recognising										
XX a protein encoded by a marker gene; and (7) a DNA chip for testing for										
XX bronchial asthma or a chronic obstructive pulmonary disease, on which a										
XX probe has been immobilised to assay a marker gene. (I) has respiratory										
XX and antiasthmatic activities, and can be used in gene therapy. The method										
XX is useful for testing for or screening for a therapeutic agent for										
XX bronchial asthma or chronic obstructive pulmonary disease. The present										
XX sequence is used in the exemplification of the present invention.										
XX Sequence 516 AA;										
XX Query Match										
XX Best Local Similarity										
XX 98.2%; Score 2682; DB 8; Length 516;										
XX 99.2%; Pred. No. 5.9e-193;										

Testing for bronchial asthma or chronic obstructive pulmonary disease by comparing the expression level of a marker gene in a biological sample from a subject with the expression level of the gene in a sample from a

PT healthy subject.

Example 11; SEQ ID NO 665; 241pp; English.

The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

Sequence 516 AA;

Query Match 98.2%; Score 2682; DB 8; Length 516;

Best Local Similarity 99.2%; Pred. No. 5.9e-193;

Matches 493; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

7 RAAPRVLFGEWLLGESSGCEGLQWLDEARTCFRVPWKHFARKDISEADARIFKAWAVA 66
20 RAAPRVLFGEWLLGESSGCEGLQWLDEARTCFRVPWKHFARKDISEADARIFKAWAVA 79
67 RGRWPPSSRGSGGPPPPAEATAERAGWNTNFCALRSTRFRVWMLDNDGDPADPHKVYALSR 126
80 RGRWPPSSRGSGGPPPPAEATAERAGWNTNFCALRSTRFRVWMLDNDGDPADPHKVYALSR 139
127 ELCWREGPGTDQTEAEAPAAVPPQGGPPGFPLATHAGLQAPGPLPAPAGDGDLLQ 186
140 ELCWREGPGTDQTEAEAPAAVPPQGGPPGFPLATHAGLQAPGPLPAPAGDGDLLQ 199
187 VQOSCLADHLLTASWADVPVTKAPGEGQGLPLTCACAGGPGLPAGELYGNVETTPSP 246
200 VQOSCLADHLLTASWADVPVTKAPGEGQGLPLTCACAGGPGLPAGELYGNVETTPSP 259
247 GPQPAALTGTEAAAPSPHQAEPYLSFSPSACTAVQSPGALDVTIMYKRTVLQKVVG 306
260 GPQPAALTGTEAAAPSPHQAEPYLSFSPSACTAVQSPGALDVTIMYKRTVLQKVVG 319
307 HPSCFTLYGPPDPVAVRATPDQVAFPSPAELPDQKQLRYTELLRHVAFGLHLELGPQL 366
320 HPSCFTLYGPPDPVAVRATPDQVAFPSPAELPDQKQLRYTELLRHVAFGLHLELGPQL 379
367 WARMGCKYVWVGPPGASSTPACLLPRNCPTPIFDVRFVFPOLVEFRARQRGSP 426
380 WARMGCKYVWVGPPGASSTPACLLPRNCPTPIFDVRFVFPOLVEFRARQRGSP 439
427 RYTIYLGFGDLSAGRPKESLVLVKLEPWLVRHLEGTQREGVSSLDSSDLCLSSAN 486
440 RYTIYLGFGDLSAGRPKESLVLVKLEPWLVRHLEGTQREGVSSLDSSDLCLSSAN 499
487 SLYDDIECFLEMELEQPA 503

Db 500 SLYDDIECFLEMELEQPA 516

RESULT 9

ADL83070

ID ADL83070 standard; protein; 516 AA.

XX ADL83070;

DT 17-JUN-2004 (first entry)

DE Human PRO83673, SEQ ID 272.

XX Immunosuppressive; Cytostatic; Antiarthritic; Antirheumatic; Antianemic;
KW Antiallergic; Muscular; Neuroprotective; Nephrotropic; Antiinflammatory;
KW Gene Therapy; PRO; B cell related disorder; cancer;
KW immune-mediated inflammatory disease; human.

OS Homo sapiens.

PN WO2004024097-A2.

PD 25-MAR-2004.

PF 15-SEP-2003; 2003WO-US029097.

PR 16-SEP-2002; 2002US-0411392P.

PA (GETH) GENENTECH INC.

PI Chiu H, Clark H, Dennis K, Fong S, Schoenfeld JR, Wood W;

PI Wu TD;

DR WPI: 2004-329389/30.

DR N-PSDB; ADL83069.

XX New PRO polypeptide, useful for diagnosing and treating a B cell related disorder, e.g. Burkitt's lymphoma, rheumatoid arthritis, autoimmune mediated hemolytic anemia, myasthenia gravis or ankylosing spondylitis.

PS Claim 10; Fig 272; 695pp; English.

XX The present invention relates to PRO proteins and their coding sequences. The PRO proteins are useful for diagnosing and treating a B cell related disorder, e.g. X-linked infantile hypogammaglobulinemia, polyaccharide antigen unresponsiveness, selective IgA deficiency, selective IgM deficiency, selective deficiency of IgG subclasses, immunodeficiency with hyper IGM, transient hypogammaglobulinemia of infancy, Burkitt's lymphoma, intermediate lymphoma, follicular lymphoma, type II hypersensitivity, rheumatoid arthritis, autoimmune mediated hemolytic anaemia, myasthenia gravis, hypoadrenocorticism, glomerulonephritis, or ankylosing spondylitis. The PRO proteins are also useful for preparing a medicament for treating a condition that is responsive to the PRO protein, e.g. cancer or immune-mediated inflammatory diseases. The PRO coding sequences are useful as hybridization probes in chromosome and gene mapping, in preparing PRO proteins, or in generating transgenic animals or knockout animals, which in turn are useful in the development and screening of therapeutically useful reagents.

SQ Sequence 516 AA;

Query Match 98.2%; Score 2682; DB 8; Length 516;

Best Local Similarity 99.2%; Pred. No. 5.9e-193;

Matches 493; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 RAAPRVLFGEWLLGESSGCEGLQWLDEARTCFRVPWKHFARKDISEADARIFKAWAVA 66

Db 20 RAAPRVLFGEWLLGESSGCEGLQWLDEARTCFRVPWKHFARKDISEADARIFKAWAVA 79

QY 67 RGRWPPSSRGSGGPPPPAEATAERAGWNTNFCALRSTRFRVWMLDNDGDPADPHKVYALSR 126

Db 80 RGRWPPSSRGSGGPPPPAEATAERAGWNTNFCALRSTRFRVWMLDNDGDPADPHKVYALSR 139

Qy	127	ELCWREGPGTDTQTEAEAPAAVPPPPGCGPPGPF ^{LA} HTHAGLQAPG ^{PL} PAPAGDKGDL ^{LL} LQA	186
Db	140	ELCWREGPGTDTQTEAEAPAAVPPPPGCGPPGPF ^{LA} HTHAGLQAPG ^{PL} PAPAGDEGDL ^{LL} LQA	199
Qy	187	VQOSCLADHLLTASWCAQDPVPTKAPGEQEG ^{LP} LTGACAGG ^{PL} PAGELYGNAVET ^{TP} SP	246
Db	200	VQOSCLADHLLTASWCAQDPVPTKAPGEQEG ^{LP} LTGACAGG ^{PL} PAGELYGNAVET ^{TP} SP	259
Qy	247	GPQPAALT ^{TT} GEAAAPESPHQAE ^{PY} LS ^{PS} PSACTAVQ ^{PS} PGALDVT ^{IM} YKGR ^{TV} LQK ^W VG	306
Db	260	GPQPAALT ^{TT} GEAAAPESPHQAE ^{PY} LS ^{PS} PSACTAVQ ^{PS} PGALDVT ^{IM} YKGR ^{TV} LQK ^W VG	319
Qy	307	HPSC ^{TF} LYGPPDPAVRATD ^{PO} QVAF ^{PS} PAEL ^{PD} QK ^{LR} YTEELLRHVAPG ^{LH} ELRG ^{PQ} L	366
Db	320	HPSC ^{TF} LYGPPDPAVRATD ^{PO} QVAF ^{PS} PAEL ^{PD} QK ^{LR} YTEELLRHVAPG ^{LH} ELRG ^{PQ} L	379
Qy	367	WARRMGKCKYVVEVGPPGSGSP ^{SP} ACLL ^{PR} NC ^{DT} PI ^{FD} FRV ^{FF} QELV ^{EF} FRAR ^{QR} RGSP	426
Db	380	WARRMGKCKYVVEVGPPGSGSP ^{SP} ACLL ^{PR} NC ^{DT} PI ^{FD} FRV ^{FF} QELV ^{EF} FRAR ^{QR} RGSP	439
Qy	427	RYTYILYGFQDLSAGRPKE ^{KS} LV ^{LK} LPW ^{LC} R ^V HEG ^{TQ} REGV ^{SS} LD ^{SS} DL ^{DL} CL ^{SS} AN	486
Db	440	RYTYILYGFQDLSAGRPKE ^{KS} LV ^{LK} LPW ^{LC} R ^V HEG ^{TQ} REGV ^{SS} LD ^{SS} DL ^{DL} CL ^{SS} AN	499
Qy	487	SLYDDIECF ^{LM} ELQPA	503
Db	500	SLYDDIECF ^{LM} ELQPA	516
RESULT 10			
ADN04327	ID	ADN04327 standard; protein; 516 AA.	
AC	ADN04327;		
XX	01-JUL-2004	(first entry)	
XX	Antipsoriatic protein sequence #358.		
DE	antipsoriatic; gene therapy; psoriasis; diagnosis.		
XX	Homo sapiens.		
XX	WO2004028479-A2.		
XX	08-APR-2004.		
XX	25-SEP-2003; 2003WO-US030907.		
XX	25-SEP-2002; 2002US-0414006P.		
PR	(GETH) GENENTECH INC.		
PA	Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WT;		
PI	Wu TD;		
XX	WPI; 2004-305105/28.		
DR	N-PSDB; ADN04326.		
XX	New PRO nucleic acid or polypeptide, useful for preparing a		
PT	pharmaceutical composition for diagnosing or treating psoriasis in a		
PT	mammal.		
XX	Claim 9; SEQ ID NO 721; 3069pp; English.		
XX	The invention relates to novel polynucleotide and polypeptides for		
CC	treating psoriasis or a sequence having at least 80% identity to the		
CC	above sequences. The nucleic acid is useful for preparing a composition		
CC	for diagnosing or treating psoriasis in a mammal. This sequence		
CC	corresponds to one of the polypeptides of the invention.		
XX	Sequence 516 AA;		
XX	SEQ		

Query Match	98.2%; Score 2682; DB 8; Length 516;
Best Local Similarity	99.2%; Pred. No. 5.9e-193;
Matches	493; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY	7	RAAPRVLFGEWLLGETISSGCEGLQWLDEARTCFRVPWKHFARKDLSEADARLFKAWAVA	66
DB	20	RAAPRVLFGEWLLGETISSGCEGLQWLDEARTCFRVPWKHFARKDLSEADARLFKAWAVA	79
QY	67	RGWRPSSRGSGGPPPEAETAERAGWKTNFRCALRRFRVMLRDNSGDPADPHKVVALSR	126
DB	80	RGWRPSSRGSGGPPPEAETAERAGWKTNFRCALRRFRVMLRDNSGDPADPHKVVALSR	139
QY	127	ELCWREGPGTDQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGPIAPAGDKGDLILQA	186
DB	140	ELCWREGPGTDQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGPIAPAGDKGDLILQA	199
QY	187	VQOSCLADHLLTASWGADPVPTKAPGEGGEGPLTGACAGGPGLPAGELYGMAVETTPSP	246
DB	200	VQOSCLADHLLTASWGADPVPTKAPGEGGEGPLTGACAGGPGLPAGELYGMAVETTPSP	259
QY	247	GPQPAALTTGEAAAPSPHQAEPLYSPSPSACTAVQEPSGALDVTIMYKGRITVLQKVG	306
DB	260	GPQPAALTTGEAAAPSPHQAEPLYSPSPSACTAVQEPSGALDVTIMYKGRITVLQKVG	319
QY	307	HPSCFTLYGPPDPAVATDPQVAFSPSAPLPPQKQLRYTEELLRHVAPGLHLELRGQPL	366
DB	320	HPSCFTLYGPPDPAVATDPQVAFSPSAPLPPQKQLRYTEELLRHVAPGLHLELRGQPL	379
QY	367	WARRMGCKVYWEVGGPPGGSASPSPTACLLPRNCDTPIPDFRVFFOELVEFRARQRRGSP	426
DB	380	WARRMGCKVYWEVGGPPGGSASPSPTACLLPRNCDTPIPDFRVFFOELVEFRARQRRGSP	439
QY	427	RYTYILGFGQDLSAGRPKEKSLVLVKLEPWLCRVHLEGTQREGVSSLDSSDDLCLSSAN	486
DB	440	RYTYILGFGQDLSAGRPKEKSLVLVKLEPWLCRVHLEGTQREGVSSLDSSDDLCLSSAN	499
QY	487	SLYDDIECFLMELEQPA 503	
DB	500	SLYDDIECFLMELEQPA 516	

RESULT 11
ADP24098
ID ADP24098 standard; protein; 516 AA.
AC AC
ADP24098;
XX AC
XX AC
XX AC
DT DT
18-NOV-2004 (first entry)
XX DT
DE PRO polypeptide SEQ ID NO:1276.
XX DE
XX DE
KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX KW
XX KW
OS Unidentified.
XX OS
XX OS
FN WO2004041170-A2.
XX FN
XX FN
PD 21-MAY-2004.
XX PD
XX PD
PF 30-OCT-2003; 2003WO-US034312.
XX PF
XX PF
PR 01-NOV-2002; 2002US-0423394P.
XX PR
XX PR
PA (GETH) GENENTECH INC.
XX PA
XX PA
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
FI Wu ID;
XX FI
XX FI
DR WPI; 2004-419628/39.
DR N-PSDB; ADP24097.
DR DR

XX New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 7; SEQ ID NO 1276; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX
SQ Sequence 516 AA;
Query Match 98.2%; Score 2682; DB 8; Length 516;
Best Local Similarity 99.2%; Pred. No. 5.9e-193;
Matches 493; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
7 RAAPRVLFGEWLLGESSGCGYGLQWLDEARTCFRVPWKHFARKOLSEADARIFKAWAVA 66
20 RAAPRVLFGEWLLGESSGCGYGLQWLDEARTCFRVPWKHFARKOLSEADARIFKAWAVA 79
67 RGRWPPSSRGSGGPPPEAETAEAGWKTNFCALRSTRFVMLRDNSSGDPADPHKVALSR 126
80 RGRWPPSSRGSGGPPPEAETAEAGWKTNFCALRSTRFVMLRDNSSGDPADPHKVALSR 139
127 ELCWREGPGTDTAEAPAAVPPQGGPGPFPLAHTHAGLQAPGLPAPAGDKGDLLLQA 186
140 ELCWREGPGTDTAEAPAAVPPQGGPGPFPLAHTHAGLQAPGLPAPAGDEGDLLLQA 199
187 VQOSCLADHLLTASWGADVPVTKAPGEGGEGPLTGACAGGPGLPAGELYGWAVETTPSP 246
200 VQOSCLADHLLTASWGADVPVTKAPGEGGEGPLTGACAGGPGLPAGELYGWAVETTPSP 259
247 GPQPAALTGTEAARAPSPHQAEPLSPSACTAVQEPSGALDVTIMYKGRVTLQKVVG 306
260 GPQPAALTGTEAARAPSPHQAEPLSPSACTAVQEPSGALDVTIMYKGRVTLQKVVG 319
307 HPSCTFLYGPDPVAVRATDPPQVAFPSPAELPDQKQLRYTEELRHVAPGLHLELRGPQL 366
320 HPSCTFLYGPDPVAVRATDPPQVAFPSPAELPDQKQLRYTEELRHVAPGLHLELRGPQL 379
367 WARMGCKVYWEVGGPPGASPSPTACLLPRNCDTPIFDPRVFFQELVEFRARQRGSP 426
380 WARMGCKVYWEVGGPPGASPSPTACLLPRNCDTPIFDPRVFFQELVEFRARQRGSP 439
427 RYTIYLGFGODLSAGRPKEKSLVLVLEPWLCHVHLEGTQREGVSSLDSSDLCLSSAN 486
440 RYTIYLGFGODLSAGRPKEKSLVLVLEPWLCHVHLEGTQREGVSSLDSSDLCLSSAN 499

QY 487 SLYDDIECFLEMELEQPA 503
DB 500 SLYDDIECFLEMELEQPA 516

RESULT 12

ADJ75411
ID ADJ75411 standard; protein; 474 AA.

XX AC ADJ75411;

XX DT 20-MAY-2004 (first entry)

XX DE Marker gene related amino acid sequence SEQ ID NO: 663.

XX KW bronchial asthma; chronic obstructive pulmonary disease;

KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.

XX OS Homo sapiens.

XX PN EP1394274-A2.

XX PD 03-MAR-2004.

XX 04-AUG-2003; 2003EP-00254857.

XX 06-AUG-2002; 2003JP-00229312.

XX 20-MAR-2003; 2003JP-00077212.

XX (GENO-) GENOX RES INC.

XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX WPI; 2004-193155/19.

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.

XX Example 11; SEQ ID NO 663; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.

```
XX SQ Sequence 474 AA;
Query Match 92.8%; Score 2533.5; DB 8; Length 474;
Best Local Similarity 93.8%; Pred. No. 7.7e-182;
Matches 472; Conservative 0; Mismatches 2; Indels 29; Gaps 1;

QY 1 MALAPERAAPRVLFGEWLLGEISSGCVGEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAAPRVLFGEWLLGEISSGCVGEGLQWLDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAWAVARGWRPPSSRGSGPPPEAETAEAGWKTNFRCALRSTRFRVWMLRNSGDPADPHK 120
DB 61 KAWAVARGWRPPSSRGSGPPPEAETAEAGWKTNFRCALRSTRFRVWMLRNSGDPADPHK 120
QY 121 VYALSRELCWREGPGTDQTEAEAPAAVPPQGGPGPFLLAHTHAGLQAPGLPAPAGDKG 180
DB 121 VYALSRELCWREGPGTDQTEAEAPAAVPPQGGPGPFLLAHTHAGLQAPGLPAPAGDKG 180
QY 181 DILLQAVQOSCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGMVAV 240
DB 181 DILLQAVQOSCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGMVAV 240
QY 241 ETTTSPGPQPAALTTGEAAAPESPHQAEPLYLSPSPSACTAVQPSGALDVTIMYKGRTV 300
DB 241 ETTTSPGPQPAALTTGEAAAPESPHQAEPLYLSPSPSACTAVQPSGALDVTIMYKGRTV 300
QY 301 LQKVGHPSCTFLYGPDPDAVRATDQQAFAFPSPAEPLDQKQLRYTEELLRHVAPGLHLE 360
DB 301 LQKVGHPSCTFLYGPDPDAVRATDQQAFAFPSPAEPLDQKQLRYTEELLRHVAPGLHLE 360
QY 361 LRGQPOLWARMGKCKVYWEVGGPPGSGASPSPTACLLPRNCDTPIFDFRVFQELVEFRAR 420
DB 361 LRGQPOLWARMGKCKVYWEVGGPPGSGASPSPTACLLPRNCDTPIFDFRVFQELVEFRAR 420
QY 392 QRRGSRYTYILGFGQDLSAGRPKEKSLVLVLEPMLCRVHLEGTQREGVSSLDSSLSL 451
DB 392 QRRGSRYTYILGFGQDLSAGRPKEKSLVLVLEPMLCRVHLEGTQREGVSSLDSSLSL 451
QY 481 CLSSANSLYDDIECFLEMELOPA 503
DB 481 CLSSANSLYDDIECFLEMELOPA 503
QY 482 CLSSANSLYDDIECFLEMELOPA 474
DB 482 CLSSANSLYDDIECFLEMELOPA 474

RESULT 13
ADJ75463
ID ADJ75463 standard; protein; 474 AA.
XX AC ADJ75463;
XX DT 20-MAY-2004 (first entry)
XX DE Marker gene related amino acid sequence SEQ ID NO:715.
XX KW bronchial asthma; chronic obstructive pulmonary disease;
XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX KW gene therapy; marker.
XX OS Homo sapiens.
XX PN EF1394274-A2.
XX PD 03-MAR-2004.
XX PF 04-AUG-2003; 2003EP-00254857.
XX PR 06-AUG-2002; 2002JP-00229312.
XX PR 20-MAR-2003; 2003JP-00077212.
XX PA (GENO-) GENOX RES INC.
XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX PI
```

```
DR XX WPI; 2004-193155/19.
PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
XX Example 11; SEQ ID NO 715; 241pp; English.
XX The present invention describes a method of testing for bronchial asthma
XX or chronic obstructive pulmonary disease. The method comprises
XX determining the expression level of a marker gene in a biological sample
XX from a subject, comparing the expression level determined with the
XX expression level of the marker gene in a biological sample from a healthy
XX subject, and judging whether the subject has bronchial asthma or chronic
XX obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX genes (S1) whose expression levels increase when respiratory epithelial
XX cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX whose expression levels decrease when respiratory epithelial cells are
XX stimulated with interleukin-13. Also described: (1) a reagent (I) for
XX testing for bronchial asthma or chronic obstructive pulmonary disease;
XX (2) a kit for screening for a candidate compound for a therapeutic agent
XX to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
XX an animal model for bronchial asthma or chronic obstructive pulmonary
XX disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX method for producing an animal model for bronchial asthma or chronic
XX obstructive pulmonary disease; (6) a therapeutic agent for bronchial
XX asthma or chronic obstructive pulmonary disease, comprising the compound,
XX a marker gene or an antisense nucleic acid corresponding to a portion of
XX the marker gene, a ribozyme, a polynucleotide that suppresses the
XX expression of the gene through an RNAi effect or an antibody recognising
XX a protein encoded by a marker gene; and (7) a DNA chip for testing for
XX bronchial asthma or a chronic obstructive pulmonary disease, on which a
XX probe has been immobilised to assay a marker gene. (I) has respiratory
XX and antasthmatic activities, and can be used in gene therapy. The method
XX is useful for testing for or screening for a therapeutic agent for
XX bronchial asthma or chronic obstructive pulmonary disease. The present
XX sequence is used in the exemplification of the present invention.
XX Sequence 474 AA;
```

Query Match 92.8%; Score 2533.5; DB 8; Length 474;
Best Local Similarity 93.8%; Pred. No. 7.7e-182;
Matches 472; Conservative 0; Mismatches 2; Indels 29; Gaps 1;

QY 421 ORRGSPRYTILGFGODLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSDL 480
DB 392 ORRGSPRYTILGFGODLSAGRPKEKSLVLKLEPWLCRVHLEGTQREGVSSLDSSSL 451
QY 481 CLSSANSLYDDIECFLEMELEQPA 503
DB 452 CLSSANSLYDDIECFLEMELEQPA 474

RESULT 14
ABM83030
ID ABM83030 standard; protein; 441 AA.
XX AC ABM83030;
XX DT 18-NOV-2004 (first entry)
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:3279.
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX OS Homo sapiens.
XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
DR WPI; 2004-329368/30.
DR N-PSDB; ACN41682.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX PS Claim 27; Page; 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for sonatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: the sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX Sequence 441 AA;

Query Match

81.5%; Score 2225.5; DB 8; Length 441;

Best Local Similarity 84.3%; Pred. No. 9.8e-159;
Matches 419; Conservative 0; Mismatches 3; Indels 75; Gaps 1;
QY 7 RAAPRVLFGEWLLGEISSCYEGLQWLDEARTCFRVPWKHFARKOLSEADARIFKAWAVA 66
DB 20 RAAPRVLFGEWLLGEISSCYEGLQWLDEARTCFRVPWKHFARKOLSEADARIFKAWAVA 79
QY 67 RGRWPPSSRGGGPPPEAETAERAGWKNTPRCALRSTRRRFVMLRDNDSGDPADPHKVYALS 126
DB 80 RGRWPPSSRGGGPPPEAETAERAGWKNTPRCALRSTRRRFVMLRDNDSGDPADPHKVYALS 139
QY 127 ELCWREGPGTQDTEAEAPAAVPPQGGPPGPFATHAGLQAPGPLPAGDKGDLLOA 186
DB 140 ELCWREGPGTQDTEAEAPAAVPPQGGPPGPFATHAGLQAPGPLPAGDKGDLLOA 163
QY 187 VQSCCLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSP 246
DB 164 -----XGCLPAGELYGWAVETTPSP 184
QY 247 GPQPAALTGEEAAAPESPHQAPYLSPPSACTAVQEPSFGALDVTIMYKGRVLOKVVG 306
DB 185 GPQPAALTGEEAAAPESPHQAPYLSPPSACTAVQEPSFGALDVTIMYKGRVLOKVVG 244
QY 307 HPSCTFLYGPDPVAVRATDPOQVAPSPAEPLDQKOLRYTELLRHVAPGLHLELGPOL 366
DB 245 HPSCTFLYGPDPVAVRATDPOQVAPSPAEPLDQKOLRYTELLRHVAPGLHLELGPOL 304
QY 367 WARRMGKCKVYWEVGGPPGSGASPTPACLLPNCCTPIEDFRVFFOELVEFRARQRGSP 426
DB 305 WARRMGKCKVYWEVGGPPGSGASPTPACLLPNCCTPIEDFRVFFOELVEFRARQRGSP 364
QY 427 RYTIYLGFGQDLSAGRPKEKSLVLKLEPWLRCVHLEGTQREGVSSLDSSDLCLSSAN 486
DB 365 RYTIYLGFGQDLSAGRPKEKSLVLKLEPWLRCVHLEGTQREGVSSLDSSDLCLSSAN 424
QY 487 SLYDDIECFLEMELEQPA 503
DB 425 SLYDDIECFLEMELEQPA 441

RESULT 15
ABM83031
ID ABM83031 standard; protein; 440 AA.
XX AC ABM83031;
XX DT 18-NOV-2004 (first entry)
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:3280.
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX OS Homo sapiens.
XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;

XX WPI: 2004-329368/30.
DR N-PSDB; ACN41683.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorders, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 440 AA;
Query Match 80.9%; Score 2210; DB 8; Length 440;
Best Local Similarity 84.1%; Pred. No. 1.4e-157;
Matches 418; Conservative 0; Mismatches 3; Indels 76; Gaps 2;
Qy 7 RAAPRVLFGEWNLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIFKAWAVA 66
Db 20 RAAPRVLFGEWNLGEISSGCGYEGQLWLDARTCFRVPWKHFARKDLSEADARIFKAWAVA 79
Qy 67 RGRWPPSSRGSGGPPPEAETAEAGWKNFRCAKSTRFVWLRDMSGDPADPHKVVALSR 126
Db 80 RGRWPPSSRGSGGPPPEAETAEAGWKNFRCAKSTRFVWLRDMSGDPADPHKVVALSR 139
Qy 127 ELCWREGFGTDQTEAEAPAAVPPPGGPGFFLAHTAGLQAPGPLPAPAGDKGDLLOA 186
Db 140 ELCWREGFGTDQTEAEAPAAVPP----- 163
Qy 187 VQOSCLADHLLTASWADPVTTKAPGEGQEGGLPLTGACAGPGPLPAGELYGWAVETTPSP 246
Db 164 -----XGPGLPAGELYGWAVETTPSP 184
Qy 247 GPQPAALTGEEAAPESPQAEPLSPSACTAVQSPSPCALDVTIMYKGRVTLOKVWG 306
Db 185 GPQPAALTGEEAAPESPQAEPLSPSACTAVQSPSPCALDVTIMYKGRVTLOKVWG 244
Qy 307 HPSCTFLYGPDPVAVRATDQQVAFPPSPAEPLDQKQLRYTEELRHVAPGLHLELRGQL 366
Db 245 HPSCTFLYGPDPVAVRATDQQVAFPPSPAEPLDQKQLRYTEELRHVAPGLHLELRGQL 304
Qy 367 WARMGKCKVYWEVGGPPGSPSTPACLLPRNCDTIFDFRVFFQELVFRAPQRGSP 426
Db 305 WARMGKCKVYWEVGGPPGSPSTPACLLPRNCDTIFDFRVFFQ-LVEFRAPQRGSP 363
Qy 427 RYTIYLGFGDLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSDLDLCSSAN 486
Db 364 RYTIYLGFGDLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSDLDLCSSAN 423
Qy 487 SLYDDIECFLEMELEQPA 503
Db 424 SLYDDIECFLEMELEQPA 440

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 16:40:43 ; Search time 44 Seconds
(without alignments)
853.375 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAAAPRVLFGEWLLG.....SANSLYDDIECFMELSQPA 503

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/ACTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2715	99.4	580	4	US-09-949-016-11716
2	2695	98.7	503	3	US-08-999-774A-10
3	2686	98.4	538	4	US-09-949-016-11715
4	2529.5	92.6	551	4	US-09-949-016-8970
5	1479	54.2	288	4	US-09-949-016-8969
6	458	16.8	427	4	US-09-949-016-6096
7	458	16.8	442	4	US-09-949-016-7374
8	454	16.6	427	3	US-08-705-771-16
9	454	16.6	427	4	US-09-417-540-16
10	376	13.8	450	2	US-08-611-280-2
11	376	13.8	450	3	US-09-195-940-2
12	376	13.8	450	3	US-09-562-466-2
13	363	13.3	486	4	US-09-949-016-7334
14	361.5	13.2	450	2	US-08-611-280-25
15	361.5	13.2	450	3	US-09-195-940-25
16	361.5	13.2	450	3	US-09-562-466-25
17	357	13.1	451	3	US-08-654-482-14
18	357	13.1	451	4	US-09-949-016-6229
19	328.5	12.0	393	3	US-09-230-371A-29
20	308.5	11.3	425	4	US-09-949-016-11483
21	306	11.2	425	3	US-09-230-371A-30
22	270.5	9.9	379	4	US-09-079-030-81
23	259.5	9.5	373	4	US-09-149-476-374
24	228	8.3	107	3	US-08-654-482-5
25	222.5	8.1	108	3	US-08-654-482-1
26	222.5	8.1	108	3	US-08-654-482-2
27	218	8.0	107	3	US-08-654-482-7

28	214	7.8	136	4	US-09-513-999C-5424	Sequence 5424, Ap
29	212	7.8	107	3	US-08-654-482-6	Sequence 6, Appli
30	198.5	7.3	102	4	US-09-513-999C-5034	Sequence 5034, Ap
31	176	6.4	325	4	US-09-949-016-6188	Sequence 6188, Ap
32	176	6.4	390	4	US-09-949-016-7930	Sequence 7930, Ap
33	169.5	6.2	802	4	US-09-252-991A-25050	Sequence 25050, A
34	167	6.1	352	4	US-09-949-016-11359	Sequence 11359, A
35	166	6.1	108	3	US-08-654-482-4	Sequence 4, Appli
36	161.5	5.9	108	3	US-08-654-482-3	Sequence 3, Appli
37	157.5	5.8	705	4	US-09-538-092-945	Sequence 945, App
38	156.5	5.7	420	4	US-09-902-540-13993	Sequence 13993, A
39	155.5	5.7	713	4	US-09-949-016-9700	Sequence 9700, Ap
40	155	5.7	561	1	US-08-642-255-52	Sequence 52, Appli
41	154.5	5.7	1064	1	US-08-642-255-62	Sequence 52, Appli
42	153.5	5.6	960	3	US-09-219-849-5	Sequence 5, Appli
43	153	5.6	349	4	US-09-919-497-76	Sequence 76, Appli
44	152	5.6	355	3	US-08-483-533-41	Sequence 41, Appli
45	152	5.6	355	3	US-09-283-471A-41	Sequence 41, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-11716
; Sequence 11716, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11716
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11716

Query Match	99.4%	Score	2715	DB	4	Length	580
Best Local Similarity	99.4%	Pred. No.	1.8e-224				
Matches	500	Conservative	1	Mismatches	2	Indels	0
Gaps	0						
QY	1	MALAPERAAAPRVLFGEWLLG	ISSGCGYEGLOWLDEARTCFRVPWKHFARKOLSEADARIF	60			
Db	78	MALAPERAAAPRVLFGEWLLG	ISSGCGYEGLOWLDEARTCFRVPWKHFARKOLSEADARIF	137			
QY	61	KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTFRFCALRSTRFVMLRDNSGDPADPHK	120				
Db	138	KAWAVARGWPPSSRGSGGPPPEAETAEAGWKTFRFCALRSTRFVMLRDNSGDPADPHK	197				
QY	121	VVALSRELWCWREGPGTDTQTEAEAPAAVPPQGGPGPFPLAHTHAGLQAPGLPAPAGDKG	180				
Db	198	VVALSRELWCWREGPGTDTQTEAEAPAAVPPQGGPGPFPLAHTHAGLQAPGLPAPAGDKG	257				
QY	181	DLLLOAVQOCSLADHLLLTASWCAADVPVTKAPGEGGGLPLTGACAGGPGCPAGELGWAV	240				
Db	258	DLLLOAVQOCSLADHLLLTASWCAADVPVTKAPGEGGGLPLTGACAGGPGCPAGELGWAV	317				
QY	241	ETTPSGPGQPAALTTGEAAAPSPHQAEPYLSFSPSACTAVQSPFGALDVTIMYKRTV	300				
Db	318	ETTPSGPGQPAALTTGEAAAPSPHQAEPYLSFSPSACTAVQSPFGALDVTIMYKRTV	377				
QY	301	LOKVGVHPSCTFLYGPDPDPVAVRATDQVAFPSPAELPDQKQLRYTEELLRHVAPGLHLE	360				

Db 378 LQKVGHPSTCTLYGPPDPVAVRATDQQAFAFSPAEPLDQKQLRYTEELLRHVAPGLHLE 437
Qy 361 LRGPOLWARRMGCKKYVWEGGPGSASPTPACLLPRNCDTPIFDFRVFFQELVEFRAR 420
Db 438 LRGPOLWARRMGCKKYVWEGGPGSASPTPACLLPRNCDTPIFDFRVFFQELVEFRAR 497
Qy 421 QRRGSPRYTYILGFGODLSAGRPKEKSLVLVLEPWLRCRVHLEGTQREGVSSLDSSDL 480
Db 498 QRRGSPRYTYILGFGODLSAGRPKEKSLVLVLEPWLRCRVHLEGTQREGVSSLDSS 557
Qy 481 CLSSANSLYDDIECFLEMELEQPA 503
Db 558 CLSSANSLYDDIECFLEMELEQPA 580

RESULT 2
US-08-999-774A-10
; Sequence 10, Application US/08999774A
; Patent No. 6274312
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Seghezzi, Wolfgang
; APPLICANT: Shanahan, Frances
; APPLICANT: Lees, Emma M.
; APPLICANT: Mcclanahan, Terrill K.
; TITLE OF INVENTION: Intracellular Regulatory Molecules;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESS: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,774A
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,818
; FILING DATE: 11-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-999-774A-10

Query Match 98.7%; Score 2695; DB 3; Length 503;
Best Local Similarity 99.0%; Pred. No. 7.6e-223;
Matches 498; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MALAPEAARVLFGEWLLGEISSGCGYEGLOWLDEARTCFRPWKHFARKDLSEADARIF 60
Db 1 MALAPEAARVLFGEWLLGEISSGCGYEGLOWLDEARTCFRPWKHFARKDLSEADARIF 60
Qy 61 KAWAVARGWPPSSRGSGPPPEAETARAGWKTNFRCLSTRRRFVMLRDNNSGDPADPHK 120

Db 61 KAWAVARGWPPSSRGSGPPPEAETARAGWKTNFRCLSTRRRFVMLRDNNSGDPADPHK 120
Qy 121 VVALSRELWCWREGPGTDDTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKG 180
Db 121 VVALSRELWCWREGPGTDDTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKG 180
Qy 181 DLIQLQAVQOSCLADHLLTASWGADPVPTKAPGEGQEGGLPLTGACAGGPGPLPAGELYGWAV 240
Db 181 DLIQLQAVQOSCLADHLLTASWGADPVPTKAPGEGQEGGLPLTGACAGGPGPLPAGELYGWAV 240
Qy 241 ETPSPGPPAALTTCGEAAAPSPHQAEPYLSPPSACTAVQPPSPGALDVTIMYKGRTV 300
Db 241 EKTSPSPGPPAALTTCGEAAAPSPHQAEPYLSPPSACTAVQPPSPGALDVTIMYKGRTV 300
Qy 301 LQKVVGHPSTCTLYGPPDPVAVRATDQQAFAFSPAEPLDQKQLRYTEELLRHVAPGLHLE 360
Db 301 LQKVVGHPSTCTLYGPPDPVAVRATDQQAFAFSPAEPLDQKQLRYTEELLRHVAPGLHLE 360
Qy 361 LRGPOLWARRMGCKKYVWEGGPGSASPTPACLLPRNCDTPIFDFRVFFQELVEFRAR 420
Db 361 LRGPOLWARRMGCKKYVWEGGPGSASPTPACLLPRNCDTPIFDFRVFFQELVEFRAR 420
Qy 421 QRRGSPRYTYILGFGODLSAGRPKEKSLVLVLEPWLRCRVHLEGTQREGVSSLDSSDL 480
Db 421 QRRGSPRYTYILGFGODLSAGRPKEKSLVLVLEPWLRCRVHLEGTQREGVSSLDSSLSL 480
Qy 481 CLSSANSLYDDIECFLEMELEQPA 503
Db 481 CLSSANSLYDDIECFLEMELEQPA 503

RESULT 3
US-09-949-016-11715
; Sequence 11715, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11715
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11715

Query Match 98.4%; Score 2686; DB 4; Length 538;
Best Local Similarity 99.4%; Pred. No. 4.9e-222;
Matches 494; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 RAAPRVLFGEWLLGEISSGCGYEGLOWLDEARTCFRPWKHFARKDLSEADARIFKAWAVA 66
Db 42 RAAPRVLFGEWLLGEISSGCGYEGLOWLDEARTCFRPWKHFARKDLSEADARIFKAWAVA 101
Qy 67 RGRWPPSSRGSGPPPEAETARAGWKTNFRCLSTRRRFVMLRDNNSGDPADPHKVYALS 126
Db 102 RGRWPPSSRGSGPPPEAETARAGWKTNFRCLSTRRRFVMLRDNNSGDPADPHKVYALS 161
Qy 127 ELWCWREGPGTDDTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKGLLQ 186
Db 162 ELWCWREGPGTDDTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKGLLQ 221
Qy 187 VQOSCLADHLLTASWGADPVPTKAPGEGQEGGLPLTGACAGGPGPLPAGELYGWAVETPSP 246

Db 222 VQSCADLLTASGADVPVTKAPGEGQLPLTGACAGGPGLPAGELYGNVAVETTPSP 281
QY 247 GPQPAALTTGEEAAAPESPHQAEPLYLSPSPSACTAVQEPSFGALDVTIMYKGRVTLQKVVG 306
Db 282 GPQPAALTTGEEAAAPESPHQAEPLYLSPSPSACTAVQEPSFGALDVTIMYKGRVTLQKVVG 341
QY 307 HPSCTFLYGPDPAPVAVRATDQQAFFSPAPLDPQKOLRYTELLRHVAPGLHLELRGPOL 366
Db 342 HPSCTFLYGPDPAPVAVRATDQQAFFSPAPLDPQKOLRYTELLRHVAPGLHLELRGPOL 401
QY 367 WARRMGCKVYVEVGGPGSPASSTPACLLPRNCPTPIEDFRVFEQELVEFRARORRGP 426
Db 402 WARRMGCKVYVEVGGPGSPASSTPACLLPRNCPTPIEDFRVFEQELVEFRARORRGP 461
QY 427 RYTIYLGFGQDLISAGRPKESLVLVKLEPWLRCRVHLEGTQREGVSSLDSSDLCLSSAN 486
Db 462 RYTIYLGFGQDLISAGRPKESLVLVKLEPWLRCRVHLEGTQREGVSSLDSSDLCLSSAN 521
QY 487 SLYDDIECFMLEQPA 503
Db 522 SLYDDIECFMLEQPA 538

RESULT 4

US-09-949-016-8970
; Sequence 8970, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8970
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8970

Query Match 92.6%; Score 2529.5; DB 4; Length 551;
Best Local Similarity 93.6%; Pred. No. 1.3e-208;
Matches 471; Conservative 1; Mismatches 2; Indels 29; Gaps 1;
QY 1 MALAPERAAPRVLFGEWLLGEISSGCVGELQWLDEARTCFRPWKHPARKDLSEADARIF 60
Db 78 MALAPERAAPRVLFGEWLLGEISSGCVGELQWLDEARTCFRPWKHPARKDLSEADARIF 137
QY 61 KAWAVARGWPSSRGSGPPPEAETARAGKTNFRCALRSTRFRVWLKNSGDPADPHK 120
Db 138 KAWAVARGWPSSRGSGPPPEAETARAGKTNFRCALRSTRFRVWLKNSGDPADPHK 197
QY 121 VVALSRELWREGPGTDQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGPLPAPAGDKG 180
Db 198 VVALSRELWREGPGTDQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGPLPAPAGDKG 257
QY 181 DILLQVQSCIALDHLLTASWAGADPVTKAPGQGLPLTGACAGGPGLPAGELYGNVAV 240
Db 258 DILLQVQSCIALDHLLTASWAGADPVTKAPGQGLPLTGACAGGPGLPAGELYGNVAV 303
QY 241 EITTPSPQPAALTTGEEAAAPESPHQAEPLYLSPSPSACTAVQEPSFGALDVTIMYKGRV 300
Db 304 -----GEEAAAPESPHQAEPLYLSPSPSACTAVQEPSFGALDVTIMYKGRV 348

QY 301 LOKVYHPSCTFLYGPDPAPVAVRATDQQAFFSPAPLDPQKOLRYTELLRHVAPGLHLE 360
Db 349 LOKVYHPSCTFLYGPDPAPVAVRATDQQAFFSPAPLDPQKOLRYTELLRHVAPGLHLE 408
QY 361 LRGPOLWARRMGCKVYVEVGGPGSPASSTPACLLPRNCPTPIEDFRVFEQELVEFRAR 420
Db 409 LRGPOLWARRMGCKVYVEVGGPGSPASSTPACLLPRNCPTPIEDFRVFEQELVEFRAR 468
QY 421 QRRGSPRYTIYLGFGQDLISAGRPKESLVLVKLEPWLRCRVHLEGTQREGVSSLDSSDL 480
Db 469 QRRGSPRYTIYLGFGQDLISAGRPKESLVLVKLEPWLRCRVHLEGTQREGVSSLDSSLSL 528
QY 481 CLSSANSLYDDIECFMLEQPA 503
Db 529 CLSSANSLYDDIECFMLEQPA 551

RESULT 5

US-09-949-016-8969
; Sequence 8969, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8969
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8969

Query Match 54.2%; Score 1479; DB 4; Length 288;
Best Local Similarity 95.2%; Pred. No. 7.3e-119;
Matches 278; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

QY 212 GEGQGLPLTGACAGGPGLPAGELYGNVAVETTPSPGQPAALTTGEEAAAPESPHQAEPLY 271
Db 1 GRGPRSCPTT-----TGPGLPAGELYGNVAVETTPSPGQPAALTTGEEAAAPESPHQAEPLY 56
QY 272 SPSPSACTAVQEPSFGALDVTIMYKGRVTLQKVGHPSCTFLYGPDPAPVAVRATDQQAFF 331
Db 57 SPSPSACTAVQEPSFGALDVTIMYKGRVTLQKVGHPSCTFLYGPDPAPVAVRATDQQAFF 116
QY 332 PSPAELPDQKOLRYTELLRHVAPGLHLELRGPOLWARRMGCKVYVEVGGPGSPASST 391
Db 117 PSPAELPDQKOLRYTELLRHVAPGLHLELRGPOLWARRMGCKVYVEVGGPGSPASST 176
QY 392 PACLLPRNCPTPIEDFRVFEQELVEFRARORRGSPPRYTIYLGFGQDLISAGRPKESLVLV 451
Db 177 PACLLPRNCPTPIEDFRVFEQELVEFRARORRGSPPRYTIYLGFGQDLISAGRPKESLVLV 236
QY 452 KLEPWLRCRVHLEGTQREGVSSLDSSDLCLSSANSLYDDIECFMLEQPA 503
Db 237 KLEPWLRCRVHLEGTQREGVSSLDSSDLCLSSANSLYDDIECFMLEQPA 288

RESULT 6

US-09-949-016-6096
; Sequence 6096, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6096
; LENGTH: 427
; TYPE: PR1
; ORGANISM: Human
; US-09-949-016-6096

Query Match      16.8%; Score 458; DB 4; Length 427;
Best Local Similarity 28.0%; Pred. No. 4.7e-31;
Matches 141; Conservative 64; Mismatches 191; Indels 108; Gaps 19;

Qy 10 PRVLFGWLGEISSGCGYGLQWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVARG 69
Db 6 PRIL--PWLVSQDLGQLEGVAVWVKSRTFRFPWKHGLRQDAQOEDFGIFQAWAETGA 63
Qy 70 WPSSRGSGPPPEAETAEAGWKTNFCALRSTRFRVWMLRDNQSDPADPHKVVALSREL 129
Db 64 YVP---GRDKPLDPT-----WKRNFSAALNRKEGLRLAEDRSKDPDHPKHIFEV--- 110
Qy 130 WREGPGTDQTEABAPAAVPPQGGPPGPFPLAHTAGLQAPGLPAPAGDKGDLLOAVQ 189
Db 111 -NSGVG---DFSQDPTSPDTNGS-----GSTSD-----TQ 136
Qy 190 SCLADHLLTASWAGADVPVTKAPGEGQGLPLTGACAGGGLPAGELYGMVAVETTPSPGP 249
Db 137 EDILDELL--GNMVLAPLP-----DPGPP-----SLAVAPEPCPQ 169
Qy 250 PAALTGEEAAPSHPQAEPLSPSPSACTAVQEP--SPGALDVTIMYKGTVLQKVGH 308
Db 170 PLRSPSLDNTPT-----FPNLGSENLKRLVPGSEWFEVTFATYRGQVFOQTISCP 223
Qy 309 SCTFLYGPDPAPVATDPQVAFPSA--ELPDQKQLRYTEELRHVAPGLHLELRGPOLW 367
Db 224 EGLLVG--SEVGRDTPGWPVTLDPGMSLTDGVMVSVRHVLSCLGGGLALWRAGQWLW 282
Qy 368 ARMGKCKVYWEVG-----GPPGSASPSTPACLLPRNCPTPIPFDFRVFFQELVEFR 418
Db 283 AORLGHCHTYVAVSEELLPSNGHGDGE-----VPKDEKGGVFDLGFIVDLITFT 333
Qy 419 AQRRGSPRYTILGFGQDLSAGRPKEKSLVVLKLEPWLCRVHLEGTQREGVSSLDSSDL 478
Db 334 EGSGR--SPRYALMFCVGSWPQDPQWTKRLVMVWVFTCLRALVEMARVGGASSENT-V 391
Qy 479 DCLSSANSLSLSDQYKAVLQDL 499
Db 392 DLHISNSHPLSLTSDQYKAVLQDL 415

RESULT 7
US-09-949-016-7374
; Sequence 7374, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7374
; LENGTH: 442
; TYPE: PR1
; ORGANISM: Human
; US-09-949-016-7374

Query Match      16.8%; Score 458; DB 4; Length 442;
Best Local Similarity 28.0%; Pred. No. 4.9e-31;
Matches 141; Conservative 64; Mismatches 191; Indels 108; Gaps 19;

Qy 10 PRVLFGWLGEISSGCGYGLQWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVARG 69
Db 21 PRIL--PWLVSQDLGQLEGVAVWVKSRTFRFPWKHGLRQDAQOEDFGIFQAWAETGA 78
Qy 70 WPSSRGSGPPPEAETAEAGWKTNFCALRSTRFRVWMLRDNQSDPADPHKVVALSREL 129
Db 79 YVP---GRDKPLDPT-----WKRNFSAALNRKEGLRLAEDRSKDPDHPKHIFEV--- 125
Qy 130 WREGPGTDQTEABAPAAVPPQGGPPGPFPLAHTAGLQAPGLPAPAGDKGDLLOAVQ 189
Db 126 -NSGVG---DFSQDPTSPDTNGS-----GSTSD-----TQ 151
Qy 190 SCLADHLLTASWAGADVPVTKAPGEGQGLPLTGACAGGGLPAGELYGMVAVETTPSPGP 249
Db 152 EDILDELL--GNMVLAPLP-----DPGPP-----SLAVAPEPCPQ 184
Qy 250 PAALTGEEAAPSHPQAEPLSPSPSACTAVQEP--SPGALDVTIMYKGTVLQKVGH 308
Db 185 PLRSPSLDNTPT-----FPNLGSENLKRLVPGSEWFEVTFATYRGQVFOQTISCP 238
Qy 309 SCTFLYGPDPAPVATDPQVAFPSA--ELPDQKQLRYTEELRHVAPGLHLELRGPOLW 367
Db 239 EGLLVG--SEVGRDTPGWPVTLDPGMSLTDGVMVSVRHVLSCLGGGLALWRAGQWLW 297
Qy 368 ARMGKCKVYWEVG-----GPPGSASPSTPACLLPRNCPTPIPFDFRVFFQELVEFR 418
Db 298 AORLGHCHTYVAVSEELLPSNGHGDGE-----VPKDEKGGVFDLGFIVDLITFT 348
Qy 419 AQRRGSPRYTILGFGQDLSAGRPKEKSLVVLKLEPWLCRVHLEGTQREGVSSLDSSDL 478
Db 349 EGSGR--SPRYALMFCVGSWPQDPQWTKRLVMVWVFTCLRALVEMARVGGASSENT-V 406
Qy 479 DCLSSANSLSLSDQYKAVLQDL 499
Db 407 DLHISNSHPLSLTSDQYKAVLQDL 430

RESULT 8
US-08-705-771-16
; Sequence 16, Application US/08705771
; Patent No. 6054289
; GENERAL INFORMATION:
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
; APPLICANT: Jian Ni and Jing-Shan Hu
; TITLE OF INVENTION: Human Genes, Sequences and
; TITLE OF INVENTION: Expression Products
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESS: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
```


QY 301 LQKVVGHP-SCFTLYGPPDPVAVRATDPOQVAPPSPAEPLDQKOLRYTELLRHVAPGLHL 359
Db 264 KELTTSPGCRISHG---HTYDVSNLDQVLFYP---DDNQKRNIEKLLSHLBERGLVL 317
QY 360 ELRGPOLWARRMKCKVYWEVGPPGSGAS--PSTPACLLPRNCDTPIFDPRVFFQELVEF 417
Db 318 WMAPDGLYAKLCSRIYWD--GPLALCSDRNK-----LERDOTCKLFTQQLSELQV 371
QY 418 RARQRGSPRYTYLGFQDLSAGRPKEKSLVLVLEPWLCR 459
Db 372 -AAHGRPAPRFQVTLFCGEEFPDPQ-RQRKLTAHVPELLAR 411

RESULT 12
US-09-562-466-2
; Sequence 2, Application US/09562466
; Patent No. 6369202
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; Grossman, Alex
; Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/562,466
APPLICATION NUMBER: 09/195,940
FILING DATE: 01-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/195,940
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-338A
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-562-466-2
Query Match 13.8%; Score 376; DB 3; Length 450;
Best Local Similarity 28.1%; Pred. No. 5.4e-24;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;
QY 16 EWLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIPKAVAVGRWPPSS 74
Db 26 QWLIDQIDSGKYPGLVWENEKSVFRIPWKHAGKQDYNEEDAALFKAWLFKGF---- 81
QY 75 RGGGPPPEAETAERAGKTNFCALASTRRFVMLRDSNG-DPADPHKVVALSRELWCWREG 133
Db 82 REGIDKPDPT-----WKTLRCALNKNSNDFEELVERSQLDISDPYKVVRI----- 127
QY 134 PGTDQTEABAPAAVPPPGGPPGPPFLAHTHAGLQAPGP-LPAPAGDKGDLILLQAVQOS 190
Db 128 -----VPEGAKKAKQLTLDQTMANGHPYPTAP---YGSUPAQOVHNY 169
QY 191 CLADHLLTASWGADPVPTKAPGEGQGLPLTGACG-----GPGLP-----AGELYGWAVE 241

Db 170 WMPPH--DRSW-RDYAPDQSHPEIYQCVTTGPRGHHWQSGCENGCOVTTGYACAPP 226
QY 242 TTPSPG-PQPAALTTEAAAPESPHQAEPYLSFSPSACTAVQSPSGALDVTIMYKGRV 300
Db 227 ESQAFGPIEPSIRSAEALA-----LSDC-----RLHICLYYRDILV 263
QY 301 LQKVVGHP-SCFTLYGPPDPVAVRATDPOQVAPPSPAEPLDQKOLRYTELLRHVAPGLHL 359
Db 264 KELTTSPGCRISHG---HTYDVSNLDQVLFYP---DDNQKRNIEKLLSHLBERGLVL 317
QY 360 ELRGPOLWARRMKCKVYWEVGPPGSGAS--PSTPACLLPRNCDTPIFDPRVFFQELVEF 417
Db 318 WMAPDGLYAKLCSRIYWD--GPLALCSDRNK-----LERDOTCKLFTQQLSELQV 371
QY 418 RARQRGSPRYTYLGFQDLSAGRPKEKSLVLVLEPWLCR 459
Db 372 -AAHGRPAPRFQVTLFCGEEFPDPQ-RQRKLTAHVPELLAR 411

RESULT 13
US-09-949-016-7334
; Sequence 7334, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7334
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7334

Query Match 13.3%; Score 363; DB 4; Length 486;
Best Local Similarity 27.5%; Pred. No. 7.9e-23;
Matches 136; Conservative 73; Mismatches 193; Indels 92; Gaps 24;
QY 16 EWLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLS-EADARIPKAVAVGRWPPSS 74
Db 61 QWLIDQIDSGKYPGLVWENEKSVFRIPWKHAGKQDYNEEDAALFKAWLFKGF---- 116
QY 75 RGGGPPPEAETAERAGKTNFCALASTRRFVMLRDSNG-DPADPHKVVALSRELWCWREG 133
Db 117 REGIDKPDPT-----WKTLRCALNKNSNDFEELVERSQLDISDPYKVVRIPEGAKK-- 169
QY 134 PGTDQTEABAPAAVPPPGGPPGPPFLAHTHAGLQAPGP-LPAPAGDKGDLILLQAVQOSCL 192
Db 170 -GAKQLTLEDPOW-----SMSHPYT-WTTPYPSLPA-----QOVHNYM 206
QY 193 ADHLTASWGADPVPTKAPGEGQGLPLT-----GACAGGGLPAGELYGWAVET 242
Db 207 PE--LDRSW-RDYVPDQHPPIYQCPMTFGPRGHHWQGPACENGQCV-TGTFYACAPPE 262
QY 243 TTPSPG-PQPAALTTEAAAPESPHQAEPYLSFSPSACTAVQSPSGALDVTIMYKGRV 301
Db 263 SOAGVPTSPSIRSAEALA-----FSDC-----RLHICLYYRIILVK 299
QY 302 QKVVGHP-SCFTLYGPPDPVAVRATDPOQVAPPSPAEPLDQKOLRYTELLRHVAPGLHLE 360
Db 300 ELTTSSPEGCRISHG---HTYDVSNLDQVLFYP---EDNGQKRNIEKLLSHLBERGVLMW 353

Search completed: May 6, 2005, 16:49:19
Job time : 47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 16:44:59 ; Search time 58 Seconds
(without alignments)
2893.107 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSYDDIECFMLELEQPA 503

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1428581 seqs, 333598853 residues

Total number of hits satisfying chosen parameters: 1428581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/prodata/1/pubpaa/FCI_NEW_PUB.pep:*
 - 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*
 - 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*
 - 6: /cgn2_6/prodata/1/pubpaa/FCIUS_PUBCOMB.pep:*
 - 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*
 - 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep:*
 - 10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep:*
 - 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep:*
 - 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
 - 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
 - 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pep:*
 - 17: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep:*
 - 18: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep:*
 - 19: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:*
 - 20: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	832	30.5	164	16	US-10-755-889-356
2	458	16.8	427	17	US-10-741-600-1291
3	458	16.8	427	17	US-10-741-600-1293
4	454	16.6	427	9	US-09-975-253-2
5	454	16.6	427	12	US-09-975-253-2
6	454	16.6	427	15	US-10-453-478-16
7	446	16.3	175	15	US-10-264-049-4045
8	431	15.8	467	17	US-10-840-1398-2
9	431	15.8	467	17	US-10-840-1398-3
10	364.5	13.3	452	17	US-10-741-600-1292
11	337.5	12.4	497	17	US-10-492-043-2
12	335.5	12.3	72	15	US-10-276-774-1410
13	332.5	12.2	473	17	US-10-492-043-30

14	328.5	12.0	393	17	US-10-492-043-5	Sequence 5, Appli
15	320	11.7	392	15	US-10-276-774-2333	Sequence 2333, Ap
16	259.5	9.5	373	10	US-09-809-391-374	Sequence 373, App
17	259.5	9.5	373	10	US-09-882-171-374	Sequence 374, App
18	259.5	9.5	373	15	US-10-164-861-374	Sequence 374, App
19	188.5	6.9	625	16	US-10-437-963-165015	Sequence 165015,
20	176	6.4	292	9	US-09-805-020-46	Sequence 46, Appl
21	176	6.4	395	15	US-10-284-049-2929	Sequence 2929, Ap
22	171	6.3	411	15	US-10-108-260A-4496	Sequence 4496, Ap
23	167	6.1	1023	10	US-09-893-519A-14	Sequence 14, Appl
24	166.5	6.1	1014	17	US-10-901-816A-12	Sequence 12, Appl
25	162	5.9	577	15	US-10-055-569A-4	Sequence 4, Appli
26	161.5	5.9	551	15	US-10-055-569A-42	Sequence 42, Appl
27	157	5.7	1413	15	US-10-288-798-24	Sequence 24, Appl
28	157	5.7	1413	15	US-10-362-892-34	Sequence 24, Appl
29	156	5.7	696	15	US-10-092-900A-272	Sequence 272, App
30	155.5	5.7	668	14	US-10-240-154-20	Sequence 20, Appl
31	155	5.7	576	15	US-10-055-569A-40	Sequence 40, Appl
32	155	5.7	576	17	US-10-898-142-13	Sequence 13, Appl
33	154.5	5.7	409	16	US-10-437-963-140753	Sequence 140753,
34	153.5	5.6	960	15	US-10-342-331-5	Sequence 5, Appli
35	153	5.6	349	9	US-09-919-497-76	Sequence 76, Appl
36	153	5.6	542	15	US-10-276-774-1386	Sequence 1386, Ap
37	152	5.6	1466	15	US-10-402-089-4	Sequence 4, Appli
38	152	5.6	1466	15	US-10-402-089-6	Sequence 6, Appli
39	152	5.6	1466	15	US-10-402-072A-4	Sequence 4, Appli
40	152	5.6	1466	15	US-10-402-072A-6	Sequence 6, Appli
41	151	5.5	369	16	US-10-437-963-162013	Sequence 162013,
42	151	5.5	507	16	US-10-437-963-187017	Sequence 187017,
43	150.5	5.5	550	9	US-09-976-740-47	Sequence 47, Appl
44	150.5	5.5	550	13	US-10-023-529-47	Sequence 47, Appl
45	150.5	5.5	550	13	US-10-023-523-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1

US-10-755-889-356
; Sequence 356, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 356
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-356

Query Match	30.5%	Score	832	DB	16	Length	164
Best Local Similarity	100.0%	Pred. No.	1e+50				
Mismatches	151	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MALAPERAPRVLFGEWLLGESSCYEGLWLDARTCFRVPWKHFARKDSEADARIF	60				
Db	1	MALAPERAPRVLFGEWLLGESSCYEGLWLDARTCFRVPWKHFARKDSEADARIF	60				
Qy	61	KAWAVARGWPPSSRGCGPPPEAETAEARAGWKTNFRFCALRSTRFRFVLMRDNSGDPADPHK	120				
Db	61	KAWAVARGWPPSSRGCGPPPEAETAEARAGWKTNFRFCALRSTRFRFVLMRDNSGDPADPHK	120				
Qy	121	VVALSRELCWREGPGTDTQTEAEAPAAVPPFQ	151				

OTHER INFORMATION: Xaa equals Ile, Leu, Phe, or Val
US-09-975-253-2

Query Match 16.6%; Score 454; DB 9; Length 427;
Best Local Similarity 28.0%; Pred. No. 1.1e-23;
Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;
10 PRVLFGEWLLGEISSCYEGLWLDARTCTPRVPMKHPARKDLSEADARIFKAWAVARGR 69
6 PRXL--PMLVSQLDLGLEGVAVWVKSRTRFRIPWKHGLRQDAQQEDFGIQAWEATGA 63
70 WPPSSRGSGPPPEAETAEAGWKTRFRALRSTRFRVMLRDNSSGDPADPHKVVALSREL 129
64 YVP---GRDKPDLPT-----WKRFNSALNRKEGLRLAEDRSKDPHDPKHYEFV--- 110
130 WREGPTDQTEAEAPAAVPPPGPGPFPLAHTAGLQAPGLPAPAGDKGDLILLQAVQQ 189
111 -NSGVG---DFSQDPTSPTNGG-----GSTSD-----TQ 136
190 SCLADHLLTASGADVPVTKAPGEGQEGPLTGACAGGPGCLPAGELYGWAVETTPSPGQ 249
137 EDILDELL--GNMVLAPLP-----DFGPP-----SLAVAPEPCPQ 169
250 PAALTGTEAAAPESPHQAEPLYSPSPSACTAVOEP--SPGALDVTIMYKGRTVLQKVVGHP 308
170 PLRSPSLNPTP-----FPNLGSENPLKRLVPGEEWEFEVTAFYRGRQVFOQTISCP 223
309 SCTFLYGGPPDPAVRATDPOQVAFPSA-ELPDQKQLRYTEELLRHVAPGLHLELRGPQLW 367
224 EGLRLVG--SEVGDRTLPGWPVTLDPGMSLTDGVMYSYVRHVLSCILGGGLALWRAGQLW 282
368 ARMGCKKYWEVG-----GPPGSASPTPACLLPRNCDTPIFDFRVFQELVEPR 418
283 AORLGHCITYWAVSEELLPSNGHGPGE-----VPKDEGGVFDLGPFIVDLITFT 333
419 ARQRGSPRYTYLFGQDLSAGRPKESLVLVLEPMLCRVHLEGTQREGVSSLDSSDL 478
334 EGSGR--SPRYALWFCVGSWPQDQPTWKRLVMKVVPCTCLRALVEMARVGGASSLENT-V 391
479 DLCLSSANSL---YDDIECFML 499
392 DLHISNSHPLSLTSDQYKAYLQDL 415

RESULT 5
US-09-975-253-2
; Sequence 2, Application US/0975253
; Publication No. US20050054033A9
; GENERAL INFORMATION:
; APPLICANT: Paul Moore et al.
; TITLE OF INVENTION: Methods and Compositions for Treating and Preventing Infection
; FILE REFERENCE: PF196P1
; CURRENT APPLICATION NUMBER: US/09/975,253
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/239,963
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals Ile, Leu, Phe, or Val
US-09-975-253-2

Query Match 16.6%; Score 454; DB 12; Length 427;
Best Local Similarity 28.0%; Pred. No. 1.1e-23;
Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;

10 PRVLFGEWLLGEISSCYEGLWLDARTCTPRVPMKHPARKDLSEADARIFKAWAVARGR 69
6 PRXL--PMLVSQLDLGLEGVAVWVKSRTRFRIPWKHGLRQDAQQEDFGIQAWEATGA 63
70 WPPSSRGSGPPPEAETAEAGWKTRFRALRSTRFRVMLRDNSSGDPADPHKVVALSREL 129
64 YVP---GRDKPDLPT-----WKRFNSALNRKEGLRLAEDRSKDPHDPKHYEFV--- 110
130 WREGPTDQTEAEAPAAVPPPGPGPFPLAHTAGLQAPGLPAPAGDKGDLILLQAVQQ 189
111 -NSGVG---DFSQDPTSPTNGG-----GSTSD-----TQ 136
190 SCLADHLLTASGADVPVTKAPGEGQEGPLTGACAGGPGCLPAGELYGWAVETTPSPGQ 249
137 EDILDELL--GNMVLAPLP-----DFGPP-----SLAVAPEPCPQ 169
250 PAALTGTEAAAPESPHQAEPLYSPSPSACTAVOEP--SPGALDVTIMYKGRTVLQKVVGHP 308
170 PLRSPSLNPTP-----FPNLGSENPLKRLVPGEEWEFEVTAFYRGRQVFOQTISCP 223
309 SCTFLYGGPPDPAVRATDPOQVAFPSA-ELPDQKQLRYTEELLRHVAPGLHLELRGPQLW 367
224 EGLRLVG--SEVGDRTLPGWPVTLDPGMSLTDGVMYSYVRHVLSCILGGGLALWRAGQLW 282
368 ARMGCKKYWEVG-----GPPGSASPTPACLLPRNCDTPIFDFRVFQELVEPR 418
283 AORLGHCITYWAVSEELLPSNGHGPGE-----VPKDEGGVFDLGPFIVDLITFT 333
419 ARQRGSPRYTYLFGQDLSAGRPKESLVLVLEPMLCRVHLEGTQREGVSSLDSSDL 478
334 EGSGR--SPRYALWFCVGSWPQDQPTWKRLVMKVVPCTCLRALVEMARVGGASSLENT-V 391
479 DLCLSSANSL---YDDIECFML 499
392 DLHISNSHPLSLTSDQYKAYLQDL 415

RESULT 6
US-10-453-478-16
; Sequence 16, Application US/10453478
; Publication No. US20030208043A1
; GENERAL INFORMATION:
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
; Jian Ni and Jing-Shan Hu
; TITLE OF INVENTION: Human Genes, Sequences and
; Expression Products
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/453,478
; FILING DATE: 04-Jun-2003
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,771
; FILING DATE: August 30, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-346 (PF196)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700

TELEFAX: 973-994-1744
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 427 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-453-478-16

Query Match 16.6%; Score 454; DB 15; Length 427;
Best Local Similarity 28.0%; Pred. No. 1.1e-23;
Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;
QY 10 PVLFGWLGESSGCVGLQWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVARG 69
DB 6 PRLX--PWLVSQDLQGLGVAWVNSRTRFRPWKHLRQDAQQDFGIFQWASATGA 63
QY 70 WPPSSRGSGPPPEAETAEAGWKTNFRCALRSTRFRVWLRDNGDPADPHKVALSRELC 129
DB 64 YVP-----GRDKPDLPT-----WKRFESALNRKGLRLAEDRSKDPHPKIYEFV---- 110
QY 130 WRGPGTDTQTEAARAPAVPPQGGPGPFPLATHAGLQAPGLPAPAGDKGDLQLQAVQQ 189
DB 111 -NSGVG----DFSQPDTSPDTNGG-----GSTSD-----TQ 136
QY 190 SCLADHLLTASWGADVPVTRKAPGEGQGLPLTGACAGGPGLPAGELYGMVAETTPSPGPQ 249
DB 137 EDLDEL--GNVLAPLP-----DPGPP-----SLVAPEPCPQ 169
QY 250 PAALTGEEAAPSPHOAEPLYSPSPSACTAVOEP--SPGALDVTIMYKGRVTLQKVVGHP 308
DB 170 PLRSPSLDNTP-----FNLGSENPXKLLVPGEEFEFTAFYRGQVQQITSCP 223
QY 309 SCTFLYGPDPDAVRATDQOQVAPSPA-ELPDOKQLRYTEELRHVAPGLHLEIRGPQLW 367
DB 224 EGLRLVG-SEVGDRTLPGWVTLPPDQMSLTDRGVMSYVRHVLSCGLGGGLALWRAGQLW 282
QY 368 ARWGKCKVYWEVG-----GPPGSASPSIPACLLPNCDDTPIFDRVFQELVEER 418
DB 283 AQLRGHCHTTWAVSEELPNSHGPDGE-----VPMKKEGVDFDLGPFVIDLITFT 333
QY 419 ARQRGSPRTYIYLGFGQDLASGRPKESLVKLEPWLRCRVHLEGTQREGVSSLSDDL 478
DB 334 EGSR--SPRYALMFCVGSWPQDQWTKRLVMVKVPTCLRALVEMARVGGASSENT-V 391
QY 479 DLCLSSANSL---YDDIECFIMEL 499
DB 392 DLHISNHPLSLTSDQYKAYLQDL 415

RESULT 7
US-10-264-049-4045
Sequence 4045, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA133PI
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: Patent In Ver. 3.1
SEQ ID NO 4045
LENGTH: 175
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE

LOCATION: (4)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (34)_FEATURE
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (90)_FEATURE
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (106)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (166)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4045
Query Match 16.3%; Score 446; DB 15; Length 175;
Best Local Similarity 55.1%; Pred. No. 1.4e-23;
Matches 102; Conservative 9; Mismatches 16; Indels 58; Gaps 8;
QY 334 PAELPDQKQLRYTEELRHVAPGLHLEIRGPQLWARRMGKCKVYWEVGGPPGSASPSTA 393
DB 2 PAXLPDQKQLRYTEELRHVAPGLHLEIRGPQLWARRMGKCKVYWEVGGPPGSASPSTA 61
QY 394 CLLPNCDDTPIFDRVFQ-----ELVEPRARQ 421
DB 62 CLLPNCDDTPIFDRVFQVQVPLAWGKPGSLGKHLPCQVLXXKSPSELVEF---- 117
QY 422 RGSPRY-TIY-----LGFGQDLASGRPKK-----SLVLVLEPWLRCRVHLEGTQREG 469
DB 118 RHGAPWLPITYLHLPGLRAG--FVSLGGPRRRPGCEAGTLVV-----PMHLEGTXXK- 167
QY 470 VSSLD 474
DB 168 VGFLD 172

RESULT 8
US-10-840-139B-2
Sequence 2, Application US/10840139B
Publication No. US2005008985A1
GENERAL INFORMATION:
APPLICANT: Schutte, Brian C.
APPLICANT: Murray, Jeffrey C.
APPLICANT: Kondo, Shinji
APPLICANT: Dixon, Michael J.
TITLE OF INVENTION: IRF6 POLYMORPHISMS ASSOCIATED WITH CLEFT LIP AND/OR PALATE
FILE REFERENCE: P06215US01
CURRENT APPLICATION NUMBER: US/10/840,139B
CURRENT FILING DATE: 2004-05-06
PRIOR APPLICATION NUMBER: US 60/468,191
PRIOR FILING DATE: 2003-05-06
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent in version 3.3
SEQ ID NO 2
LENGTH: 467
TYPE: PRT
ORGANISM: Homo sapiens
US-10-840-139B-2

Query Match 15.8%; Score 431; DB 17; Length 467;
Best Local Similarity 27.7%; Pred. No. 4.9e-22;
Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 20;
QY 1 MALAPERAAPRVLFGWLLGEISSGCVGLQWLDEARTCFRVPWKHFAR-KDLSEADARI 59
DB 1 MALHPR-----RVRLKPLWLVAVQVDSGLYPGLIWLHRSKRFOIPWKATRHSPQEEENTI 56
QY 60 FKAWAVARGWRPPSSRGSGPPPEAETAEAGWKTNFRCALRSTRFRVWLRDNGSD-PADP 118

```
Db 57 FKAWVETGKY---QGVDDDPAPK-----WKAQLRCALNKSRFENLMDYGTKEVPMNP 107
Qy 119 HKVYALSRCLWREGPGTQTEAEAPAAVPPQGGPPGFLAHTAGLQAPGLPAPAGD 178
Db 108 VKIY-----QVC-----DIPQGG-----SIINPGSTGSAPWD 135
Qy 179 KGDLLQAVQOCLADHLHTAGWADPVFTKAPGEGGLPL-----TGACAGGPGLPAGE 234
Db 136 EXD---NDVDEDEDE--LDQSHHVPIQDTFFFLNINGSMPAPASVGNCSVGNCSPEAV 191
Qy 235 LYGWAVETTPSPQPOPAALTTCGAAPESPHQAEPVLSPPSACTAVQEPSGALDVTIM 294
Db 192 ---W-----PKTEPLEMEVPO--APIQFFYSSPELWISSLPMT-----DLDIKFP 231
Qy 295 YKGRTVLQKV-VGHP--SCTFLYGPDP-----AVRATDPQVAPFPAELPDQQLRYT 346
Db 232 YRGKEYGQMTVSNPGCRLFGDGLGMPDQBELFGPVSLQGVKPPGPHITNEKQLFT 291
Qy 347 BELLRHVAFLHLELRLGPGQWARRMKCKVYVEVGGPPGSGASPTPACILLPRNCOTPIPD 406
Db 292 SKLLDVMDRGLILEVSHGAIYAIRLCQCKVYWSGPCAPSLVAPN---LIERQKKVKLPC 347
Qy 407 FRVFGQELVEFRARQRGRSPRTIYLGFQODLSAGRPKESLVLYKLEPWLGRVHLEGTQ 466
Db 348 LETFLSDLIAHQKGQIEKQPPFEIYLCFGEWPDGKPLERKILLVQVIVPVARMIYEMPS 407
Qy 467 REGVSLDSDLDLCLSSANSLYDDIECFMEL 499
Db 408 GDFTRSFDSGVRQLQISTPD-IKDNIVAQLKQL 439

RESULT 9
US-10-840-139B-3
; Sequence 3, Application US/10840139B
; Publication No. US20050089885A1
; GENERAL INFORMATION:
; APPLICANT: Schutte, Brian C.
; APPLICANT: Murrat, Jeffrey C.
; APPLICANT: Kondo, Shinji
; APPLICANT: Dixon, Michael J.
; TITLE OF INVENTION: IRF6 POLYMORPHISMS ASSOCIATED WITH CLEFT LIP AND/OR PALATE
; FILE REFERENCE: P06215US01
; CURRENT APPLICATION NUMBER: US/10/840,139B
; CURRENT FILING DATE: 2004-05-06
; PRIOR APPLICATION NUMBER: US 60/468,191
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 3
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: 5' UTR to Met
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(1)
; OTHER INFORMATION: Metille
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (2)..(2)
; OTHER INFORMATION: Ala2Val
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (6)..(6)
; OTHER INFORMATION: Arg6fs
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (17)..(17)
; OTHER INFORMATION: Gln17fs
; FEATURE:
```

```
; NAME/KEY: VARIANT
; LOCATION: (18)..(18)
; OTHER INFORMATION: Val18Met
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (18)..(18)
; OTHER INFORMATION: Val18Ala
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (39)..(39)
; OTHER INFORMATION: Pro39Ala
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (60)..(60)
; OTHER INFORMATION: Trp60Gly
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (61)..(61)
; OTHER INFORMATION: Ala61Gly
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (66)..(66)
; OTHER INFORMATION: Lys66Thr
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (68)..(68)
; OTHER INFORMATION: Gln68X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (70)..(70)
; OTHER INFORMATION: Gly70Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (76)..(76)
; OTHER INFORMATION: Pro76Ser
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (82)..(82)
; OTHER INFORMATION: Gln82Lys
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (84)..(84)
; OTHER INFORMATION: Arg84Cys
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (84)..(84)
; OTHER INFORMATION: Arg84His
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (89)..(89)
; OTHER INFORMATION: Lys89Glu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (90)..(90)
; OTHER INFORMATION: Ser90Gly
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (92)..(92)
; OTHER INFORMATION: Glu92X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (98)..(98)
; OTHER INFORMATION: Asp98His
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (118)..(118)
; OTHER INFORMATION: Gln118X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (156)..(156)
; OTHER INFORMATION: His156fs
; FEATURE:
; NAME/KEY: VARIANT
```

```
; LOCATION: (186)..(186)
; OTHER INFORMATION: Cys186X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (388)..(388)
; OTHER INFORMATION: Lys388Glu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (393)..(393)
; OTHER INFORMATION: Gln393X
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (412)..(412)
; OTHER INFORMATION: Arg412X
; FEATURE:
; NAME/KEY: VARIANT

Query Match 15.8%; Score 431; DB 17; Length 467;
Best Local Similarity 27.7%; Pred. No. 4.9e-22;
Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 20;

Qy 1 MALAPERAPRVLFGEWLLGEISSGCGEGLWLDDEARTCFRPWKHFAR-KDLSEADARI 59
Db 1 MALHPR----RVRLKPWLVAQVDSGLYPGLIWLHRDSKRFQIPWKATRHSPQOESENTI 56
Qy 60 FKAWAVARGRPSPSSRGGGPPPEAETAERAGWKTNFRCALRSTRFRVMLRDNISGD-PADP 118
Db 57 FKAWAVETGKY---QEGVDDPDPAK-----WKAQLRCALNKSRFENLMDYDGTKEVPMNP 107
Qy 119 HKVYALSRELWCWREGPGTQTEAEAPAAVPPPGGPGPFLAHTHAGLQAPGLPAPAGD 178
Db 108 VKIY---QVC-----DIPQPG-----SIINPGSTGSAPWD 135
Qy 179 KGDLLLAQVQOSCLADHLLTASMGADPVPTKAPGEGEGLPL----TGACAGGPGLPAGE 234
Db 136 EKD---NDVDEEDEDE-LDQSHHVPIDQTFPFLNINSGMAPASVGNCSVCNCSPEAV 191
Qy 235 LYGMVAVETTPSPGPQPAALTTGEAAAPESPHQAEPLYSLSPSPSACTAVQEPSGALDVTIM 294
Db 192 ---W-----PKTEPLEMEVPO--APIQPFYSSPELWISLPMT-----DLDIKFQ 231
Qy 295 YKQRTVLQKV-VGHP-SCFTLYGPPDP-----AVRATDPOQVAFPSPAELPQKQLRYT 346
Db 232 YRGKEYQGTMTVSNPQGCRLFYGDGLGMPDQBELFGVSLQEVKFGPGEHITNEKOKLFT 291
Qy 347 EELLRHVAPGLHLELPGPOLWARRMGKCKVYVGVGGPPGSGASPSPTACLLPRNCDTPIFD 406
Db 292 SKLLDVMDRGLILEVSGHAIYAIRLCQCKVYNSGPCAPSLVAPN---LIERQKVKLFC 347
Qy 407 FRVFFOELVEFRARQRRGSPRYTYIGFGODLSAGRPKEKSLVLVLEPMLCRVHLEGTQ 466
Db 348 LETFLSDLIAHQKGQIEKQPPFEIYLCFGEWPDGKPLERKILVQVIVVARMYEMFS 407
Qy 467 REGVSSILDSSDLLCLSSANSLYDDIECFMEL 499
Db 408 GDFTRFDSGVSRLQISTPD-IKDNIVAQLKQL 439

RESULT 10
US-10-741-600-1292
; Sequence 1292, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1292
; LENGTH: 452
; TYPE: PRT
```


Qy	122	YALSRELCWREGPGTDQTEAEAPAAVPPQGGPPGFFLAHTHAGLQAPGFLPAPAGDKGD	181
Db	66	YEFV-----NSGVG-----DFSQDTSPTDNGG-----GSTSD	93
Qy	182	LLLQAVQQSCLADHLLTASWGADPVPTKAPGEGQGLPUTGACAGGPGLPAGELYGWAWE	241
Db	94	-----TQEDIIDELL-GNNVLAPLP-----DPGPP-----SLA	120
Qy	242	TTTSPGQPQAALTGTGAAPESPHQAEPYLSPPSACTAVQBP-SFGALDVTIMYKRTV	300
Db	121	VAPEPCQPLRSPSLDNPTP-----FPNLGSPENPLKELLVPGEBEVEVTAFYGRQV	174
Qy	301	LQKVUGHPSCTFLYGGPDPVAVRATDQOVAFPSPA-ELPDQQLRYTEBELLRHVAPGLHL	359
Db	175	FOQTISCPEGLRVG-SEVGDRTLPGWPVTLFDPGMSLTDRGVMYVRHVLSCLGGLAL	233
Qy	360	ELRGPOLWARRMKCKVYWEVG-----GPPGSASPTPACLLPRNCDTPIPDFRVF	410
Db	234	WRAGQWLMAQRLLCHTYWAVSEELLPNSGHGDGE-----VPKDKEGGVFDLGP	284
Qy	411	F-----QELVEFRARQRGSPRYTIYLGFGODLSAGRPKKSLLVLEPWLCRVHLEG	464
Db	285	IVGSLGPPDLITFTEGSGR-SPRYALMFCVGSWPQDQWTKRLVMVKVPTCLRALVEM	343
Qy	465	TORGVSLLDSSDLCLSSANSL-----YDDIECFMEL	499
Db	344	ARVGGAASSLENT-VDLHISNSHPLSLTSDQYKAYLQDL	380

Search completed: May 6, 2005, 16:50:52
Job time : 60 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 16:28:18 ; Search time 41 Seconds
(without alignments)
1180.415 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731
Sequence: 1 MALAPERAAPRVLFGEWLLG.....SANSLYDDIECFLEMELEQPA 503

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	780	28.6	491	2 S56753	interferon regulat
2	482.5	17.7	504	2 G02474	interferon regulat
3	440	16.1	459	2 JC6520	interferon regulat
4	376	13.8	450	2 S57837	lymphoid-specific
5	333.5	12.2	424	2 A35861	interferon consensus
6	328.5	12.0	393	2 A45017	transcription fact
7	306	11.2	425	2 A45064	interferon consensus
8	259.5	9.5	399	2 JC4592	transcription fact
9	224	8.2	203	2 S57836	lymphoid-specific
10	183	6.7	1334	2 T50568	probable multi-dom
11	176	6.4	325	2 I52998	interferon regulat
12	176	6.4	1146	2 A35887	collagen, cornea-s
13	167	6.1	349	2 A53340	interferon regulat
14	166	6.1	325	2 B31595	interferon regulat
15	166	6.1	329	2 A31595	interferon regulat
16	165.5	6.1	328	2 A36330	interferon regulat
17	160.5	5.9	1400	2 T31555	hypothetical prote
18	160	5.9	1747	2 A45974	collagen alpha 1(X
19	159.5	5.8	416	1 SKXIAG	dermal gland prote
20	159	5.8	576	2 T36729	probable serine/th
21	158.5	5.8	839	2 T04859	extensin homolog F
22	157.5	5.8	705	2 A35363	synapsin I splice
23	156.5	5.7	1464	2 S59856	collagen alpha 1(I
24	156	5.7	381	2 T27806	hypothetical prote
25	156	5.7	1857	2 S31212	collagen alpha 1(X
26	156	5.7	1888	2 S78476	collagen alpha 1(X
27	156	5.7	3149	1 QQB88	BPLF1 protein - hu
28	155	5.7	801	2 T29018	hypothetical prote
29	154	5.6	660	1 QQB83	BHLF1 protein - hu

30	154	5.6	3530	2 A59266	unconventional myo
31	150.5	5.5	691	2 A25704	synapsin I - rat
32	150	5.5	704	2 A30411	synapsin Ia - rat
33	149.5	5.5	1952	2 T48814	hypothetical prote
34	148.5	5.4	560	1 WPHUM	mullerian inhibiti
35	148	5.4	384	2 H70580	hypothetical prote
36	148	5.4	1464	1 CGHU1S	collagen alpha 1(I
37	148	5.4	1466	1 CGHU7L	collagen alpha 1(I
38	147	5.4	301	2 T21314	hypothetical prote
39	147	5.4	633	2 B40983	collagen alpha 1(X
40	146.5	5.4	1042	1 CGCH1S	collagen alpha 1(I
41	145	5.3	296	2 A31219	collagen 1 - Caeno
42	145	5.3	1207	2 T00378	XIRAA0641 protein -
43	145	5.3	2715	2 T13049	eyelid - fruit fly
44	144.5	5.3	1315	2 A56101	collagen alpha 1(X
45	144.5	5.3	1546	1 CGHU2E	collagen alpha 2(X

ALIGNMENTS

RESULT 1

S56753
interferon regulatory factor 3 - chicken
C:Species: Gallus gallus (Chicken)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S56753
R:Grant, C.E.; Vasa, M.Z.; Deeley, R.G.
Nucleic Acids Res. 23, 2137-2146, 1995
A:Title: CIRP-3, a new member of the interferon regulatory factor (IRF) family that is re
A:Reference number: S56753; MUID:9534365; PMID:7541908
A:Accession: S56753
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-491 <GRA>
A:Cross-references: UNIPROT:Q90643; EMBL:U20338; NID:G790580; PIDN:AAA86995.1; PID:G790580
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 28.6%; Score 780; DB 2; Length 491;
Best Local Similarity 38.5%; Pred. No. 9.8e-40;
Matches 208; Conservative 60; Mismatches 180; Indels 92; Gaps 17;

QY	2	ALAPERAAPRVLFGEWLLG	ETSSGCGYEGLOWLDEARTCFRVWKGHPARKDLSEADARIK	61
DB	3	ALDSEGDAQKLRFGPWL	NAVSSGLYRGLCWIDPDRIRIPRIWKINARKDVTSSDVEIEK	62
QY	62	AWAVAGRWPPSSRGSGG	PPPEAETAEACAGKTNFRCALSTRFVWLRDNSGDPADPHKV	121
DB	63	AWAKASGRY	-----EGNAEDPAKWKTNFRCALRSTHFMLEDKNCNDDPHKV	111
QY	122	VALSRELWCWREGPGT	DTQTEAEPAAVPPQ--GGP-----PGPFLAHTHAGLQ----	167
DB	112	YAVA	-----SGVPNDRSGSGPVAGALQQPQLLNHDLALENTPT	152
QY	168	-APGFLPAPAGDKGL	-LLQAVQOCLADHLITASGADPVPTKAPGEG--QEGPLPTG-	222
DB	153	DSTEGVAAALTVDL	DLQLQSVLQHCNISAL-----GSQPTLWHTGDALPEDALLPGQ	207
QY	223	-ACAGGPGLPAGELYGW	-AVETTPSPGQP-----AALTTEGAAAPESHQAEF	269
DB	208	DGCLFPG	-----QFQDWRQLEPFLLLGNQPLTGGCGQDGAGALPVSECALPAPSPAE	262
QY	270	YL	-----SPSPSACTAVQPSGALDVTIMYKGRITVLQKVVGHPSCFTFLYGGPDDPAVRATD	325
DB	263	LLFQANPAPP	PAGDIGGLPPLLIITYRKVNYQEVDSDSRCLVAQLDPAV--AE	320
QY	326	PQQVAFPPSAELPD	QKQLRYTEELRHVAPGHLRLRGQLWARRMGKCKVYNEV-----	380
DB	321	QRLVLFPPSPASLP	DPDRRYTDLLE--VAGLRLEQRAGQLLATRLKCKVFWALSQQLE	378
QY	381	GGPGSASPS	TACILLPRNCDTPIFDPRVFFQELVEFRARQGRGSPRYTIYLGFQGDLSA	440
DB	379	GGEP	-----PLNLLHRDQETIIFDFRVFCTELDRDSDRSRSPDFTIFLCFGQCFSS	431

A>Description: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon

A:Reference number: S57838

A:Accession: S57838

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-450 <MAT>

A:Cross-references: EMBL:U11692; NID:g790199; PIDN:AAA75309.1; PID:g790200

R:Eisenbeis, C.F.; Singh, H.; Storb, U.

Genes Dev. 9, 1377-1387, 1995

A:Title: PIP, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcrip

A:Reference number: 149359; MUID:95317607; PMID:7797077

A:Accession: 149359

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-450 <RES>

A:Cross-references: EMBL:U34307; NID:g976446; PIDN:AAA75283.1; PID:g976447

C:Genetics:

A:Introns: 72/3; 135/1; 164/3; 213/1; 249/1; 367/1; 404/3

C:Superfamily: lymphoid-specific interferon regulatory factor

C:Keywords: alternative splicing

Query Match 13.8%; Score 376; DB 2; Length 450;
Best Local Similarity 28.1%; Pred No 1.5e-15;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

Qy 16 EWLLGEISSGCGYGLWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVARGRWPPSS 74
Db :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
26 QWLDIQDSGKYPGVLWNEEKSVFRIPWKHAGQDYNEEDAALFKAWLFKGKF---- 81
Qy 75 RGGGPPPEAETAERAGKWNTRFCALSTRFRVMDNSG-DPADPHKVYALSRELCLWREG 133
Db :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
82 REGIDKKDPPT-----WKTRLRCALKNSDFEELVERSQLDISDPYKVYRI----- 127
Qy 134 PGTDQTAEAPAAVPPPOGGPPGP---PLATHAGLAQGPLPAPAGDKGLLIQAQQOS 190
Db :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
128 -----VPEGAKGAKQLTDDTOTAMGHYPWTAP----YGSLPAQQQHNY 169
Qy 191 CLADHLTLTASMGADPVPTKAPGEGEGELPLTCACAG---GPGLP-----AGELYGWAVE 241
Db :|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:
170 MWPHH--DRSW-RDYAPDOSHEIPIQCPTVFGPRHHHQGPCSCNGCQVTGTFYACAPP 226
Qy 242 TTPSPG-PQPAALTGTEAAAPESPHQAEPLYSPSACTAVQEPSGALDVIMYKGRVT 300
Db :|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:
227 ESQAPGIPIEPSIRSAEA-----LSDC-----RLHICLYRDILV 263
Qy 301 LOKVVGHP-SCTFLYGGPDPAVRATDPOOVAFPSPAELPDQOLRYTEELLRHVAPGLHL 359
Db :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:
264 KELTTTSPEGCRISHG---HTYVSNLDQVLPFPYP---DDNGQRKNIEKLSSHLELGLVL 317
Qy 360 ELRGPOLWARMGCKVKVVEYGGPPGSGAS--PSTPACILLPNCDTFFDRFVFFQELVEF 417
Db :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:
318 WMAPDGLYAKLCSRIYWD--GPLALCSDRNK-----LERDOTCKLFDTOQFISELVQVF 371
Qy 418 RARQRGRSPRTVIYLGFOQLDSAGRKPKEKSLVLKLEPWLCR 459
Db :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:
372 -AHGRPARPRFQVTLCFGEFFPDQ-QRKLTITAHVEPLLAR 411

RESULT 5

A35861

interferon consensus sequence-binding protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 09-Jul-2004

C:Accession: A35861

R:Driggers, P.H.; Emnist, D.L.; Gleason, S.L.; Mak, W.H.; Marks, M.S.; Levi, B.Z.; Flanc

Proc. Natl. Acad. Sci. U.S.A. 87, 3743-3747, 1990

A:Title: An interferon gamma-regulated protein that binds the interferon-inducible enhancer

A:Reference number: A35861; MUID:90251633; PMID:2111015

A:Accession: A35861

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-424 <DRI>

A:Cross-references: UNIPROT:P23611; GB:M32489; NID:g194088; PIDN:AAA317878.1; PID:g309320

```
QY 76 GGGPPPEAETAEAGWKTFRCALRSTRFVLMRLDMSG-DPADPHKYVALSRELQWREGP 134
Db 70 -----KEGDTGGPAVWKTRLCALNKSSEFKEVPERGMDVAEPYKYQLLPGLVSGQP 124
QY 135 GTDQTEAABAAVPPPOGGPPGLFATHAGLQAPGLPAPAGDKGLLQAVQOOSCLAD 194
Db 125 GTQKVPKSG-----RQHSVS-----SERKEED-----AMQNCTLSP 156
QY 195 HLLTASGADVPVTKAPGEGQGLPLTGACAGGGLPAGELYGWAVETTPSPGPAAALT 254
Db 157 SVLQDSLNNEE-----EGASG-----GAVHSDIG-----SSSSSSPPFQVETDT 196
QY 255 TGEAAAPESPHQAE-----PYLSPSPSACTAQVEPSG-ALDVTIMYKGRVLOKVGHP 309
Db 197 T-----EAPFGDQDSLEFLP-----PEPYSLLLTIFYNGRVVGAQVQSLD 240
QY 310 CTFLYGPPDPVAVRATDPOQVAFPSPAELPDQKQLRYTELLRHVAPGLHLELRGPQW 369
Db 241 CRLV---AEPGSGSSMEQVLPKPGPLEP-----TORLLSQLERGILVASNPRGLFVQ 291
QY 370 RMGKCKVYVEGGPPGSGASSTPACILPRNCDTPIPDFRVFFQELVEFRARQRRG-SPRY 428
Db 292 RLCEPIISWAPQAPPGGPH-----LLPSNECEVELFETAYFCRDLV--RYFQGLGPPPKF 345
QY 429 TIVLFGQDLSAGRPKEKSLVVKLEPWLCKRVHLEGT--OREGVSSL 473
Db 346 QVTINFWEESHGSHHTQNLITVWQEAFAFYLLLEQTPEQQAILLS 392

RESULT 7
A45064
interferon consensus sequence binding protein, ICSBP (DNA-binding domain) - human
C;Species: Homo sapiens (man)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1998
C;Accession: A45064
J. Weisz, A.; Marx, P.; Sharf, R.; Appella, E.; Driggers, P.H.; Ozato, K.; Levi, B.Z.
J. Biol. Chem. 267, 25589-25596, 1992
A;Title: Human interferon consensus sequence binding protein is a negative regulator of
A;Reference number: A45064; MUID:93094284; PMID:1460054
A;Accession: A45064
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-425 <WEI>
A;Experimental source: lung and blood
A;Note: sequence extracted from NCBI backbone (NCBI:P120312)
C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; transcription regulation

Query Match 11.2%; Score 306; DB 2; Length 425;
Best Local Similarity 23.4%; Pred. No. 2.3e-11;
Matches 121; Conservative 67; Mismatches 174; Indels 156; Gaps 21;

QY 16 EWLGETSSCYEGLOWLDEARTCFRVPWKHFARKDLSE-BADARIFKAWAVARGWRPSS 74
Db 12 QWLIEQIDSSMYPGLIWENEESKMFRIPWKHAGQYDQVNEVDASIFKAWAVFKGK 67
QY 75 RGGPPPEAETAEAGWKTFRCALRSTRFVLMRLDMSG-DPADPHKYVALSREL 128
Db 68 -----KEGDKAEPATWKTRLCALNKSDFEEVTDRLSDIDSEPYKYRVIVPEEQKCK 121
QY 129 -----CWREGPQDQTEAABAAVPPPOGGPPGLFATHAGLQAPGLPAPAGDKGD 181
Db 122 LGVATACVNEVTEMEGRSEIDELIKEPS-----VDDYMGIMKRSPPDA----- 168
QY 182 LLLQAVOOSCLADHLLTASGADVPVTKAPGEGQGLPLTGACAGGGLPAGELYGWAVE 241
Db 169 -----CRSQ--LLPDW-----WAHE 181
QY 242 TTSPGPQPAALTGEEAAAPESPHQAEPLYLSPSACTAVQEPSGALDVTIMYKGRV 301
Db 182 --PSTGRR--LVTGYTTY-DAAHSAP-----SQWISFYGGKLVG 217
```

```
QY 302 QKVVGHP-SC-----TFLYGPPDPVAVRATDPOQVAFPSPAELPDQKQLRYTEEL 349
Db 218 QATTTCPEGCRSLSLQGPLGPKLYGP-----EGLELVRFPPADTIPSERQOVTRKL 270
QY 350 LRIVAFG-LHLEURGQOLWARRMGKCKVYVEGGPPGSGASSTPACILPRNCDTPIPDFR 408
Db 271 FGHLERGVLHSSRQG--VFVXRLCQGRVFCSGNAVVCGRPNK-----LERDEVVQVFDTS 324
QY 409 VFQELVEFRARORRGSPRYTIVLFGQDLSAGRPKEKSLVVKLEPWLCKRVHLEGTQRE 468
Db 325 QFFRELQQQFVNSQGR-LPDGRVVLCEGEEFPDMPAPLRSKLILVQIE-----QLYVRQLAEE 379
QY 469 GVSSDLSDDL-----DLCLSSANSLY 489
Db 380 AGKSCGAGSVMQAPEEPFPDQVFRMEFPCASHQORSFF 417

RESULT 8
JC4592
transcription factor ISGF3 gamma chain - mouse
N;Alternate names: interferon-stimulated gene factor 3 gamma chain
C;Species: Mus musculus (house mouse)
C;Date: 10-Apr-1996 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004
C;Accession: JC4592; S71599
R;Suhara, W.; Yoneyama, M.; Yonekawa, H.; Fujita, T.
J. Biochem. 119, 231-234, 1996
A;Title: Structure of mouse interferon stimulated gene factor 3 gamma (ISGF3 gamma/p48)
A;Reference number: JC4592; MUID:97037063; PMID:8882710
A;Accession: JC4592
A;Molecule type: mRNA
A;Residues: 1-399 <SUH>
A;Cross-references: UNIPROT:Q61179; EMBL:U51992; NID:g1263309; PIDN:AAC52494.1; PID:g1361;
A;Experimental source: L929 cells
R;Kawakami, T.; Matsumoto, M.; Sato, M.; Harada, H.; Taniguchi, T.; Kitagawa, M.
FEBS Lett. 358, 225-229, 1995
A;Title: Possible involvement of the transcription factor ISGF3gamma in virus-induced ex
A;Reference number: S71599; MUID:95145714; PMID:7843405
A;Accession: S71599
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-87 <KAW>
C;Genetics:
A;Gene: Isgf3g
A;Map position: 14
C;Function:
A;Description: responsible for specific interaction with the promoter element, interferon
ed gene factor 3, the primary regulator of type I interferon responses; involved in type
A;Note: induced by interferon-alpha and interferon-beta
C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; signal transduction; transcription factor
P;10-117/Domain: DNA binding #status predicted <DNB>

Query Match 9.5%; Score 259.5; DB 2; Length 399;
Best Local Similarity 22.5%; Pred. No. 1.3e-08;
Matches 105; Conservative 64; Mismatches 192; Indels 105; Gaps 14;

QY 17 WLLGETSSCYEGLOWLDEARTCFRVPWKHFARKDLSE-ADARIFKAWAVARGWRPSSR 75
Db 15 WIVEQESGHGFFGVCVDDAAKTFRIPWKHAGQDFREDQDAAFKAWALFKEK----- 68
QY 76 GGGPPPEAETAEAGWKTFRCALRSTRFVLMRLDMSG-DPADPHKYV-----ALSREL 128
Db 69 -----HKDGDIGHPAVWKTRLCALNKSSEFEEVPERGMDVAEPYKYRILPAGTLPNQ 124
QY 129 CWREGPQDQTEAABAAVPPPOGGPPGLFATHAGLQAPGLPAPAGDKGLLQAVQ 188
Db 125 RNQSPCKRSISCVSPERENMENGRTNGVNHSDSGSNIGG-----GGNGS----- 171
QY 189 QSLADHLLTASGADVPVTKAPGEGQGLPLTGACAGGGLPAGELYGWAVETTPSPGP 248
Db 172 -----NRSDSNCSNCELEEG-----AG----- 189
QY 249 QPAALTGEEAAAPESPHQAEPLYLSPSACTAVQEPSGALDVTIMYKGRVTLQKVGH 308
```

```
Db 190 -----TTEATIREDFVLEHQLPLN-----SDYSLLTFTYIGRNVGKTVHSL 233
Qy 309 SCTFLYGGPDPAVRATDPOQVAFPPGPAELPDQKQLRYTEELRHVAPGLHLELPGPQLWA 368
Db 234 DCRUVAERSD---SESSMEQVEFPK-----DPLEPTQHLLNQLDRGVLVASNSRGLEFV 284
Qy 369 RMGKCKVYVGGPPGSPGASPTACLLPRNCDTPIFRVFBELVFRARQRG-SPR 427
Db 285 QRLCPIPIISWNAPEAPPQPGPH---LLPSNKCVELFKTTFYCRDLAQV---PQGGPPPK 338
Qy 428 YTIYLGQDLSAGRPKESLVLVKLEPWLCHVHLEGTQREGVSSL 473
Db 339 FOATLHFEESPGSHSQENLITVOMEQAFARHLEKTIPEEKAAL 384

RESULT 9
S57836
Lymphoid-specific interferon regulator factor - mouse (fragments)
C:Species: Mus musculus (house mouse)
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S57836
R:Matsumura, T.; Grossman, A.; Mittrucker, H.W.; Siderovski, D.P.; Kiefer, F.; Kawakami
Nucleic Acids Res. 23, 2127-2136, 1995
A>Title: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon regula
A:Reference number: S57836; MUID:95334364; PMID:7541907
A:Accession: S57836
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-203 <MAT>
A:Cross-references: EMBL:U11692
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 8.2%; Score 224; DB 2; Length 203;
Best Local Similarity 34.8%; Pred. No. 8.7e-07;
Matches 57; Conservative 22; Mismatches 49; Indels 36; Gaps 6;

Qy 16 EWLGEISSCYEGQLWLDEARTCFRVPWKHPARKDLS-EADARIFKAWAVARGRWPSS 74
Db 4 QMLIDQISGKYPGLWENEKSVFRIPWKHAGQDYNEEDALFKAWLFKGF---- 59
Qy 75 RGGGPPPEAETAEAGKTNFRCAIARSTRFRVWLRDNGS-DEADPHKVALSRELCWREG 133
Db 60 REGIDKPDPT-----WKLRLCALNKSNDFFELVERSOLDISDPYKVYRI----- 105
Qy 134 PGTDQTEAEAPAAVPPQGGPGPFLAHAGI---QAPGLPA 174
Db 106 -----VPEGQRKNIEKLLSHLERGLVLMWAPDGLYA 136

RESULT 10
T50568
Probable multi-domain regulatory protein [imported] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T50568
R:Redenbach, M.; Kieser, H.M.; Denapaita, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A>Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
A:Reference number: Z20556; MUID:97000351; PMID:8843436
A:Accession: T50568
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1334 <RED>
A:Cross-references: UNIPROT:Q9RKB9; EMBL:AL133220; PIDN:CAB61705.1
A:Experimental source: strain A3(2)
C:Genetics:
A>Note: SCC75A.05c

Query Match 6.7%; Score 183; DB 2; Length 1334;
Best Local Similarity 28.1%; Pred. No. 0.0018;
Matches 106; Conservative 25; Mismatches 150; Indels 96; Gaps 22;
```

```
Qy 18 LLGEISSCYEGQLWLDEARTCFRVPWKHPARKDLS-EADARIFKAWAVARGRWPSSRGG 77
Db 146 VLADLPRTAEAAARW---ETRHFEALRARHTAALDGLQAEHSLPELTALCDGH----- 195
Qy 78 GPPPEAETAEAR-----AGWKNFRCALRSRRFRVWLRDNGS-DPA-----D 117
Db 196 -PLDEPQALRLRALRDSGRTAELAAAYEAVRR--LLADRLGTDPGPELRLHAELLSPS 252
Qy 118 PHKVYALSRELCWREGP-----CTDQTEAEAPAAVPPQGGP-PGPFLAHAGI 166
Db 253 FTPTPGRSRTPGWTSPPGSPASGAGAACTDVASGAGAAAGDPDPASGPASGPAVAPGSGGG 312
Qy 167 QAPGLPAPAGDKGDLLOAVQOCLADHLITASWGADPVPTKABCEQEGELPLTGACAG 226
Db 313 PAPGWPAPGTPAGSSTAPPDHTASAAD---TA-----PAPGPTSAFGTA---PAAGTAAP 362
Qy 227 GPGL--PA-GELY--GWA--VETTPSPGPQPAALTTGEA-----AAPESPHQAEPYL 271
Db 363 APGTAGPAGTSYAPGTAPVAGTTTAPGTAPGTAPGTAGPARDTSYAPGTAPVAGTTTAPGT 422
Qy 272 SPSPSACTAV-----QEPSFGALDVTIMYKGRVTLQKVGVHPSCTFLYGPDPAPVRAT 324
Db 423 APAGSTPAGTPVAPGTAPAGPQPA---DGR---RPVTGPASGT---GPG---AAT 468
Qy 325 DPQQVA-----FPSPA 335
Db 469 PPEAAAAAAGSAPSPA 485

RESULT 11
I52998
interferon regulatory factor 1 - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I52998
R:Cha, Y.; Sims, S.H.; Ronine, M.F.; Kaufmann, M.; Deisseroth, A.B.
DNA Cell Biol. 11, 60S-611, 1992
A>Title: Human interferon regulatory factor 1: intron/exon organization.
A:Reference number: I52998; MUID:93000481; PMID:1382447
A:Accession: I52998
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <RES>
A:Cross-references: UNIPROT:P10914; GB:L05072; NID:G184648; PIDN:AAA36043.1; PID:G184649
C:Genetics:
A:Gene: GDB:IRF1
A:Cross-references: GDB:127269; OMIM:147575
A:Map position: 5q31.1-5q31.1
A:Introns: 29/3; 63/1; 122/1; 138/3; 182/1; 223/1; 239/3; 285/1

Query Match 6.4%; Score 176; DB 2; Length 325;
Best Local Similarity 23.6%; Pred. No. 0.0011;
Matches 69; Conservative 35; Mismatches 108; Indels 80; Gaps 12;

Qy 11 RVLFGEWLLGEISSCYEGLOWLDEARTCFRVPWKHPARK--DLSEADARIFKAWAVARG 68
Db 5 RMRMRPWLQEMQINSQIPLGIWINKEEMIIFQIPWKHAAKHGWINDK-DACLERSWAIHTG 63
Qy 69 RWPPSSRGGPPPEAETAEAGKTNFRCAIARSTRFRVWLRD---NSGDPADPHKVYALS 125
Db 64 RY-----KAGEKEPDPT-----WKNFRCAIARSTRFRVWLRD---VRVIRML 112
Qy 126 RELC---WREGPGTDQTEAEAPAAVPPQGGPPGPF-----LAHTHAGIQAQPLPA 174
Db 113 PPLTNQKQKRSKSRDAKSKAKKSCGDSPTDFSDGLSSSTLPDDHSSYTVFG----- 168
Qy 175 PAGDKGDLLOAVQOCLADHLITASWGADPVPTKAPGEGQEGELPLTGACAGGGLP---- 231
Db 169 -----YMQDLEVEQALT-----PALSPCAVSSTLPDWH 196
Qy 232 -----AGELYGWAVETTPSPGPQPAALTTGEAAAPESPHQAEPYLSPS 274
Db 197 IPVEVVPDSTSDLYNFQV-----SPMPTSEATTDEEGKLPEDINKLLEQS 244
```

RESULT 12

A38587 collagen, cornea-specific - chicken
C;Species: Gallus gallus (chicken)
C;Date: 14-Feb-1992 #sequence_revision 15-Aug-1997 #text_change 09-Jul-2004
C;Accession: S16501; A38587
R;Marchant, J.K.; Linsemayer, T.F.; Gordon, M.K.
Proc. Natl. Acad. Sci. U.S.A. 88, 1560-1564, 1991
A;Title: cDNA analysis predicts a cornea-specific collagen.
A;Reference number: A38587; MUID:91142213; PMID:1705041
A;Accession: S16501
A;Molecule type: mRNA
A;Residues: 1-1145 <NAR>
A;Cross-references: UNIPROT:Q90584; EMBL:M60172; NID:g211609; PIDN:AAA48703.1; PID:g2116
A;Accession: A38587
A;Molecule type: mRNA
A;Residues: 1-174, 'X', 176-233, 'X', 235-344, 'X', 346-408, 'X', 410-499, 'X', 501-876, 'X', 878-11
A;Cross-references: GB:M60172
C;Keywords: cornea

Query Match 6.4%; Score 176; DB 2; Length 1146;
Best Local Similarity 23.8%; Pred. No. 0.004;
Matches 109; Conservative 31; Mismatches 132; Indels 186; Gaps 26;
QY 66 ARGWRPSSRG-----GPPPEAETAEARAGKTNFRCALRSTRFRVMLRDSGDPAD 117
DB 348 ARGPPGSGDTGPGLTGPGQPGCLPGNCPGPGAK-----GEPGA 387
QY 118 PHKYALSLRCLWREGTDTQAEAPAAVPPGPGPFLAHTAGLAQ- GP--LP- 173
DB 388 PGKVIS-----AEGSTIALPGPP-GPPGPIGTGPGVPGVPGAGLPG 431
QY 174 --APACDKG--DILLOAV--QSCADHLLTASWG--ADPVTKAPGEQGLPLTGAC 224
DB 432 QQGRGKSGAVEVVIETIKTEVSSLASQMLSDQGRAGPPGPPGESVQGLP---GP 488
QY 225 AGGPGPLPAGLYGWAVETTPS--PGQPAALTGCE--AAAPESP-----HQAEPYLS 272
DB 489 RGPPLPG-----PSGPPRGSSVSSEITFVSPPGPPGPPKPGQGE----- 533
QY 273 PSSACTAVQSPGALDVTIMYKRTVLQKVGHPSCTFLYGPDPVAVRATDP----- 326
DB 534 PGRGFTG--EPGEPGLPGSSHGGTVMQGPFP-----GPPGPKDAGVGPAGPIP 585
QY 327 -----QQV-----AFSPALPDOKQLRYTEELL-----RHVAPGLHLELRGP 364
DB 586 GTSRGSRQIQGPPGPPGPPGPPGPGSSQEIQQVYADYKSDNVHRHLYTG----- 637
QY 365 QLWARRMGKCKVYVEVGGPPGSGASPSPPACLLPRNCPTPIFD----- 407
DB 638 -----VQSPFG-----TTADGKNFDFAEALATRVMSYVTSDDH 678
QY 408 -----RVFQELVFRARQ-----RGSP 426
DB 679 YQSFASSVSTTSVLYQELNMQREIRQYLVGRGPP 716

RESULT 13

A53340 interferon regulatory factor 2 - human
N;Alternate names: transcription repressor IRF2
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A53340; S06894; A32828
R;Chia, Y.; Deisseroth, A.B.
J. Biol. Chem. 269, 5279-5287, 1994
A;Title: Human interferon regulatory factor 2 gene. Intron-exon organization and function.
A;Reference number: A53340; MUID:94148994; PMID:8106512
A;Accession: A53340
A;Status: not compared with conceptual translation
A;Molecule type: DNA

A;Residues: 1-349 <CHA>
A;Cross-references: UNIPROT:PI4316; GB:L24442
R;Itoh, S.; Harada, H.; Fujita, T.; Mimura, T.; Taniguchi, T.
Nucleic Acids Res. 17, 8372, 1989
A;Title: Sequence of a cDNA coding for human IRF-2
A;Reference number: S06894; MUID:90045964; PMID:2813069
A;Accession: S06894
A;Molecule type: mRNA
A;Residues: 1-57, 'R', 59-349 <ITO>
A;Cross-references: EMBL:X15949; NID:g33966; PIDN:CAA34073.1; PID:g33967
R;Harada, H.; Fujita, T.; Miyamoto, M.; Kimura, Y.; Maruyama, M.; Furia, A.; Miyata, T.;
Cell 58, 729-739, 1989
A;Title: Structurally similar but functionally distinct factors, IRF-1 and IRF-2, bind to
A;Reference number: A32828; MUID:89354547; PMID:2475256
A;Accession: A32828
A;Molecule type: mRNA
A;Residues: 1-68, 'I', 70-96, 'R', 98-130, 'ER', 133-152, 'GF', 155-163, 'A', 165-188, 'D', 190-210, 'A',
8, 'T', 310-313, 'PAP', 318-319, 'TP', 322-349 <HAR>
A;Cross-references: GB:J03168; NID:g198455; PIDN:AAA9333.1; PID:g293676
A;Experimental source: clones 2 and 5
C;Genetics:
A;Gene: GDB:IRF2
A;Cross-references: GDB:127270; OMIM:147576
A;Map position: 4q35.1-4q35.1
C;Keywords: DNA binding; transcription regulation
Query Match 6.1%; Score 167; DB 2; Length 349;
Best Local Similarity 35.0%; Pred. No. 0.004;
Matches 36; Conservative 18; Mismatches 39; Indels 10; Gaps 3;
QY 11 RVLFGEWLLGEISSGCGEGLQWLDEARTCFRVPKHFARKDLS-EADARIFKAWAVARGR 69
DB 5 RMRMPWLEEQINSNTIPGLKWLNEKKTFQIPWMAARHGWDVKDAPLFRNWAHTGK 64
QY 70 WPPSSRGSGPPPEAETAEARAGKTNFRCALRSTRFRVMLRDSN 112
DB 65 HQP-----GVDPKDPKT-----WKNFRCAWNSLPDIEEVKDKS 98
RESULT 14
B31595 interferon regulatory factor 1 - human
C;Species: Homo sapiens (man)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: B31595; S04075
R;Miyamoto, M.; Fujita, T.; Kimura, Y.; Maruyama, M.; Harada, H.; Sudo, Y.; Miyata, T.;
Cell 54, 903-913, 1988
A;Title: Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifically
A;Reference number: A90903; MUID:88311092; PMID:3409321
A;Accession: B31595
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-325 <M12>
A;Cross-references: UNIPROT:PI0914
R;Maruyama, M.; Fujita, T.; Taniguchi, T.
Nucleic Acids Res. 17, 3292, 1989
A;Title: Sequence of a cDNA coding for human IRF-1.
A;Reference number: S04075; MUID:89263736; PMID:2726461
A;Accession: S04075
A;Molecule type: mRNA
A;Residues: 1-325 <MAR>
A;Cross-references: EMBL:X14454
C;Keywords: DNA binding; transcription regulation
Query Match 6.1%; Score 166; DB 2; Length 325;
Best Local Similarity 23.4%; Pred. No. 0.0043;
Matches 67; Conservative 33; Mismatches 106; Indels 80; Gaps 12;
QY 17 WLIGETSSGCGEGLQWLDEARTCFRVPKHFARK--DLSEADARIFKAWAVARGRWPSS 74
DB 11 WLEMQINSQIPGLIWINKEEMILEIPWKAHKGHDINK-DACLFRSWAHTGRY---- 65
QY 75 RGGGPPPEAETAEARAGKTNFRCALRSTRFRVMLRD---NSGPDADPHKYALSLRELCL-- 129

```

Db      66 KAGEKEPDKT-----WKNFRCAMNSLPDIEVKDQSRNKGSSA--VRVYRMLPPLTKN 118
QY      130 -WREGPGTDQTEAEAPAAVPPQGGPGPF-----LAHTHAGLQADGPLPAPAGDKG 180
Db      119 QRKERSKSSRDAKSKAKSKSCDSSPDTFSDGLSSSTLPDDHSSYTFG-----168
QY      181 DILLQAVQOQSLADHLLTASGADPVPTKAPGEGQEGPLTGACACAGGPGLP-----231
Db      169 -----YMQDLEVEQALT-----PALSPCA VSTLPDWHIPVEV 202
QY      232 ---AGELYCWAVETTPSPQPAALTTGAAAPESPHQAEPLYLSPS 274
Db      203 PDSTSLYNFQV-----SPMPSISEATTDEEGKLPEDIMKLEQS 244

RESULT 15
A31595
interferon regulatory factor 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C:Accession: A31595
R:Miyamoto, M.; Fujita, T.; Kimura, Y.; Maruyama, M.; Harada, H.; Sudo, Y.; Miyata, T.;
Cell 54, 903-913, 1988
A:Title: Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifies
A:Reference number: A90903; MUID:88311092; PMID:3409321
A:Accession: A31595
A:Molecule type: mRNA
A:Residues: 1-329 <MY>
A:Cross-references: UNIPROT:P15314; GB:M21065; NID:g198458; PIDN:AAA39334.1; PID:g293677
C:Keywords: DNA binding; transcription regulation

Query Match      6.1%; Score 166; DB 2; Length 329;
Best Local Similarity 21.6%; Pred. No. 0.0043;
Matches 65; Conservative 40; Mismatches 88; Indels 108; Gaps 12;

QY      11 RVLFGEWLGEISSGCGYGLQWLDEARTCFRVFWKHFKARK--DLSEADARIFKANAVARG 68
Db      5  RMRMRPWLEMQINSQIPGLIWINKEEMIFQIPWKHAAGHWDINK-DACLFRSWAIHTG 63
QY      69 RWPPSSRGGPPPEAEATERRAGWKTNFRCALRSTRFVMLRD---NSGDPADPHKVYAL- 124
Db      64 RY-----KAGEKEPDKT-----WKNFRCAMNSLPDIEVKDQSRNKGSSA--VRVYRML 112
QY      125 -----SRELWCWREGPGT-----136
Db      113 PPLTRNQRKERSKSSRDTKSKTKRKLCDGVSPDTFSDGLSSSTLPDDHSSYTTQGYLQ 172
QY      137 -----DQTEAEAPAAVPPQGGPGPFPLAHTHAGLQ-----APGPLPAPA 176
Db      173 DLDMERDITPALSPCVV-----SSLSSEWHMQMDIIPDSTTDLYNLQVSPMPSTSEA 224
QY      177 GDKGDLQLQAVQOQSLADHLLTASGADPVPTKAPGEGQEGPLTGACACAGGPGLPAGELY 236
Db      225 ATDED-----EGKTAEDLMKLFQSEWQPHIDGKY-----LLNEFGTQLSSVY 270
QY      237 G 237
Db      271 G 271
```

Search completed: May 6, 2005, 16:45:40
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:42:52 ; Search time 15 Seconds
(without alignments)
1390.840 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731

Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSLYDDIECFLEMELEQPA 503

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2715	99.4	503	1 IRF7 HUMAN	Q92985 homo sapien
2	1562	57.2	457	1 IRF7 MOUSE	P70434 mus musculus
3	780	28.6	491	1 IRF3 CHICK	Q90643 gallus gall
4	487.5	17.9	498	1 IRF5 HUMAN	Q13568 homo sapien
5	458	16.8	427	1 IRF3 HUMAN	Q14653 homo sapien
6	444.5	16.3	497	1 IRF5 MOUSE	P56477 mus musculus
7	438	16.0	419	1 IRF3 MOUSE	P70671 mus musculus
8	436	16.0	467	1 IRF6 MOUSE	P97431 mus musculus
9	431	15.8	467	1 IRF6 HUMAN	Q14896 homo sapien
10	376	13.8	450	1 IRF4 MOUSE	Q64287 mus musculus
11	363	13.3	451	1 IRF4 HUMAN	Q15306 homo sapien
12	338	12.4	425	1 ICSB CHICK	Q90871 gallus gall
13	333.5	12.2	424	1 ICSB MOUSE	P23611 mus musculus
14	328.5	12.0	393	1 IRIF HUMAN	Q00978 homo sapien
15	308.5	11.3	426	1 ICSB HUMAN	Q02556 homo sapien
16	259.5	9.5	399	1 IRIF MOUSE	Q61179 mus musculus
17	176	6.4	325	1 IRF1 HUMAN	P10914 homo sapien
18	167	6.1	349	1 IRF2 HUMAN	P14316 homo sapien
19	167	6.1	1083	1 T2D3 HUMAN	Q00268 homo sapien
20	166	6.1	329	1 IRF1 MOUSE	P15314 mus musculus
21	166	6.1	349	1 IRF2 MOUSE	P23906 mus musculus
22	165.5	6.1	328	1 IRF1 RAT	P23570 rattus norv
23	164	6.0	348	1 IRF2 CHICK	Q98925 gallus gall
24	163	6.0	313	1 IRF1 CHICK	Q90876 gallus gall
25	159.5	5.8	439	1 XP2_XENLA	P17437 xenopus lae
26	157.5	5.8	705	1 SYNI HUMAN	P17600 homo sapien
27	156	5.7	1888	1 CAIE CHICK	P32018 gallus gall
28	156	5.7	3149	1 TEGU BEV	P03186 epstein-bar
29	155	5.7	670	1 SYNI MOUSE	Q88935 mus musculus
30	154	5.6	660	1 YHL1 BEV	P03181 epstein-bar
31	154	5.6	3530	1 MY15 HUMAN	Q9ukn7 homo sapien
32	152.5	5.6	1464	1 CA13 MOUSE	P08121 mus musculus
33	150	5.5	704	1 SYNI RAT	P09951 rattus norv

34	148.5	5.4	560	1 MIS HUMAN	P03971 homo sapien
35	148	5.4	1464	1 CA11 HUMAN	P02452 homo sapien
36	148	5.4	1466	1 CA13 HUMAN	P02461 homo sapien
37	146.5	5.4	1453	1 CA11 CHICK	P02457 gallus gall
38	145.5	5.3	1460	1 CA11 CANFA	Q9xej7 canis fami
39	145.5	5.3	1516	1 CA1H HUMAN	P39060 homo sapien
40	145	5.3	296	1 CC01 CAEL	P08124 caenorhabdi
41	144.5	5.3	1527	1 CA1H MOUSE	P39061 mus musculu
42	144.5	5.3	1736	1 CA2B HUMAN	P13942 homo sapien
43	144	5.3	415	1 SYNI CANFA	O62732 canis fami
44	144	5.3	503	1 CA19 CHICK	P12106 gallus gall
45	144	5.3	706	1 SYNI BOVIN	P17599 bos taurus

ALIGNMENTS

RESULT 1

IRF7_HUMAN
ID IRF7_HUMAN STANDARD; PRT: 503 AA.
AC Q92985; O00331; O00332; O00333; O75924;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interferon regulatory factor 7 (IRF-7).
GN IRF7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Spleen;
RA Grossman A., Nicholl J., Antonio L., Luethy R., Suggs S.,
RA Sutherland G.R., Mak T.W.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RX MEDLINE=97459673; PubMed=9315633;
RA Zhang L., Pagano J.S.;
RT "IRF-7, a new interferon regulatory factor associated with Epstein-Barr virus latency".
RL Mol. Cell. Biol. 17:5748-5757(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM D).
RX MEDLINE=99003279; PubMed=9786932;
RA Au W.-C., Moore P.A., Lafleur D.W., Tombal B., Pitha P.M.;
RT "Characterization of the interferon regulatory factor-7 and its potential role in the transcription activation of interferon A genes".
RL J. Biol. Chem. 273:29210-29217(1998).
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-STIMULATED RESPONSE ELEMENT (ISRE) IN IFN PROMOTERS AND IN THE Q PROMOTER (QP) OF EBV NUCLEAR ANTIGEN-1 (EBNA1).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A (SHOWN HERE), B/BETA, C/GAMMA AND D/H; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U73036; AAB17190.1; -
EMBL; U53830; AAB80686.1; -
EMBL; U53831; AAB80688.1; -
EMBL; U53832; AAB80690.1; -

```
DR EMBL; AF076494; AAC70999.1; -.
DR HSP; P23906; 21RF.
DR TRANSFAC; T04674; -.
DR Genew; HGNC.6122; IRF7.
DR MIM; 605047; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator;
KW Alternative splicing.
FT DNA BIND 13 122
FT VARSPPLIC 1 6
FT VARSPPLIC 152 164
FT VARSPPLIC 165 503
FT VARSPPLIC 228 256
FT CONFLICT 179 179
FT CONFLICT 412 412
FT CONFLICT Q -> K (IN REF. 2).
FT CONFLICT Q -> R (IN REF. 3).
SQ SEQUENCE 503 AA; 54278 MW; AA6A39E0E272727C CRC64;

Query Match 99.4%; Score 2715; DB 1; Length 503;
Best Local Similarity 99.4%; Pred. No. 3.8e-148;
Matches 500; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALAPERAPRVLFGEWLLGEISGCGYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60
DB 1 MALAPERAPRVLFGEWLLGEISGCGYEGLOWLDEARTCFRVPWKHFARKDLSEADARIF 60
QY 61 KAVAVARGRPSSRGSGGPPPEAETAEAGWKTNFRCALRSTRFRVWMLRDSGDPADPHK 120
DB 61 KAVAVARGRPSSRGSGGPPPEAETAEAGWKTNFRCALRSTRFRVWMLRDSGDPADPHK 120
QY 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKG 180
DB 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDEG 180
QY 181 DILLQAVQOQSLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
DB 181 DILLQAVQOQSLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELYGWAV 240
QY 241 ETTTSPGQPAAALTGGAAPESPQAEPLYSPPSACTAVQSPSCALDVTIMYKRTV 300
DB 241 ETTTSPGQPAAALTGGAAPESPQAEPLYSPPSACTAVQSPSCALDVTIMYKRTV 300
QY 301 LQKVVGHPSCFTFLYGPDPDPAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLHLE 360
DB 301 LQKVVGHPSCFTFLYGPDPDPAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLHLE 360
QY 361 LRGPQLWARRMGKCKVYWEVGGPPGSGASPTPACLLPRNCOTPIFDPRVFQELVEFRAR 420
DB 361 LRGPQLWARRMGKCKVYWEVGGPPGSGASPTPACLLPRNCOTPIFDPRVFQELVEFRAR 420
QY 421 QRGSPRYTYIYLGFGQDLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSDL 480
DB 421 QRGSPRYTYIYLGFGQDLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSLSL 480
QY 481 CLSSANSLYDDIECFLEMELEQPA 503
DB 481 CLSSANSLYDDIECFLEMELEQPA 503

RESULT 2
IRF7 MOUSE STANDARD; PRT; 457 AA.
ID IRF7 MOUSE
AC P70434;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interferon regulatory factor 7 (IRF-7).
```

```
GN IRF7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RA Grossman A., Nicholl J., Antonio L., Luethy R., Suggs S.,
RA Sutherland G.R., Mak T.W.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-
CC STIMULATED RESPONSE ELEMENT (ISRE) IN IFN PROMOTERS AND IN THE Q
CC PROMOTER (OP) OF EBV NUCLEAR ANTIGEN-1 (EBNA1) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U73037; AAB18626.1; -.
CC HSP; P23906; 11RF.
CC MGD; MGI:1859212; 11rf7.
CC InterPro; IPR001346; IRF.
CC Pfam; PF00605; IRF; 1.
CC PRINTS; PR00267; INTERREGFCT.
CC ProDom; PD002355; IRF; 1.
CC SMART; SM00348; IRF; 1.
CC PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator.
FT DNA BIND 11 122
FT SEQUENCE 457 AA; 51222 MW; 30B102C668F56142 CRC64;

Query Match 57.2%; Score 1562; DB 1; Length 457;
Best Local Similarity 62.6%; Pred. No. 1.7e-82;
Matches 315; Conservative 36; Mismatches 100; Indels 52; Gaps 7;

QY 3 LAPERAAPRVLFGEWLLGEISGCGYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFKA 62
DB 1 MAEVRGVQVRVLFDMGLLGEVSSQYEGLOWLNEARTVFRVPWKHFGRRDLDEEDAQIFKA 60
QY 63 WAVARGRWPPSSRGSGGPPPEAETA---ERAGWKTNFRCALRSTRFRVWMLRDSGDPADPH 119
DB 61 WAVARGRWPPSG-VNLPPPEAAERERRRRGKTNFRCALHSTGRFILRODNGSDPVDPH 119
QY 120 KYVALSRELCWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDK 179
DB 120 KYVELSRELGSTVGPAATENREEVLSNALPTQGVSGFLARENAGLQTPSPIL--LSSDA 177
QY 180 GDLLQAVQOQSLADHLLTASWGADVPVTKAPGEGQGLPLTGACAGGPGLPAGELY-GW 238
DB 178 GDLLQVLQVYS---HILESEGADVPVQPAQCEQDRV-----YEEPVAW 220
QY 239 AVETTTSPGQPAAALTGGAAPESPQAEPLYSPPSACTAVQSPSCALDVTIMYKGR 298
DB 221 QVEAVPSRPPQPALT-----ERSLGFDTIMYKGR 252
QY 299 TVLQKVVGHPSCFTFLYGPDPDPAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLH 358
DB 253 TVLQAVVGHPRCVFLYSPHAPAVRTSEPDVTFPSPAELPDQKQLHYTETLLQHVSEGLQ 312
QY 359 LELRGQLWARRMGKCKVYWEVGGPPGSGASPTPACLLPRNCOTPIFDPRVFQELVEFR 418
DB 313 LELRGPSLWALRMGKCKVYWEVGSPPGTTGTPSTPPQLLERNRHTPIFDSTFTFRELEEFR 372
QY 419 ARORRGSPRYTYIYLGFGQDLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSDL 478
DB 373 ARRRQSGPHYTYIYLGFGQDLSAGRPKESLVLVKLEPWLVRVHLEGTQREGVSSLDSSSL 432
```

```

QY 479 DCLSSANSYDDIECFLEMLQ 501
DB 433 GCLSSNSTSYEDIEHFLMDLQ 455

RESULT 3
IRF3_CHICK
ID IRF3 CHICK STANDARD; PRT; 491 AA.
AC Q90643;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Interferon regulatory factor 3 (IRF-3).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95334365; PubMed=7541908;
RA Grant C.E., Vasa M.Z., Deeley R.G.;
RT "cIRF-3, a new member of the interferon regulatory factor (IRF)
RL Nucleic Acids Res. 23:2137-2146(1995).
CC -!- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH
CC OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT
CC FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT
CC (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U20338; AAA86995.1; -
DR HSP; P15314; 11F1.
DR InterPro: IPR001346; IRF.
DR Pfam: PF00605; IRF. 1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator.
FT DNA BIND 14 112 TRYPTOPHAN PENTAD REPEAT.
FT SEQUENCE 491 AA; 54441 MW; 54441 MW; CAEC2AA8BE976D9 CRC64;

Query Match 28.6%; Score 780; DB 1; Length 491;
Best Local Similarity 38.5%; Pred. No. 6.6e-38;
Matches 208; Conservative 60; Mismatches 180; Indels 92; Gaps 17;

QY 2 ALAPARAPRVLEGEWLLGEISSGVEGLWLDIARTCFVFWKHFKARLDSEADARIFK 61
DB 3 ALDSEGAOKLREFGPWLNAVSSGLYRGLCWIDPDRRIFRIPWKHARKDVTSSDVEIFK 62
QY 62 ANAVARGRPSPSGGPPPEAETARAGWKTFRCALRSTRFVLMRDNSSGDPADPHKV 121
DB 63 AWAKASGRY-----EGNAEDPAKWTNFRCALRSTHFMFLERDKCNDDPHKV 111
QY 122 YALSRELWCVEGPTDQTEAEAPAAVPPQ--GGP-----PGPFLATHAGLQ---- 167
DB 112 YAVA-----SGVPNDRGSGVPVAGALQQQPQLLNHHDLALENTPT 152
QY 168 -APGPIAPAGDKGLD-LQAVQOCSCLADHLLTASWGADVPVTKAPGEG--QGLPLTG- 222
DB 153 DSTEGVAAAALTVQDLQLQSLVQHCHNISAL-----GSQPTLWAHTGDALPEDALLPGQ 207

```

RESULT 4

```

IRP5_HUMAN
ID IRP5 HUMAN STANDARD; PRT; 498 AA.
AC Q13568; Q9BQF0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 5 (IRF-5).
GN IRF5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grossman A., Mittrucker H.W., Lantonio L., Mak T.W.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U51127; AAA96056.1; -
DR EMBL; BC004201; AAH04201.1; -
DR EMBL; BC004139; AAH04139.1; -
DR HSP; P23906; 11RF.
DR Genew; HGNC:6120; IRF5.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA BIND 16 118 TRYPTOPHAN PENTAD REPEAT.
FT DOMAIN 142 149 POLY-GLU.
FT CONFLICT 161 175 EDVKWPTLQPTLR -> DAVQSGPHMTPYSLLKEDVKW
FT SEQUENCE 498 AA; 56044 MW; 01B2ED95C28384E8 CRC64;

```



```
QY 479 DLCLSSANSL---YDIECFMLMEL 499
DB 392 DLHISNHPLSLTSQYKAYLQDL 415

RESULT 6
IRF3_MOUSE
ID IRF3_MOUSE STANDARD; PRT; 497 AA.
AC P56477;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Interferon regulatory factor 5 (IRF-5).
GN IRF5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lymph node;
RA Grossman A., Kondo S., Antonio L., Mak T.W.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC 1- SUBCELLULAR LOCATION: Nuclear.
CC 2- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF028725; AAB81997.1; -
CC HSSP; P23906; IIRF.
CC MGD; MGI:1350924; IRF5.
CC InterPro; IPR001346; IRF.
CC Pfam; PF00605; IRF. 1.
CC PRINTS; PR00267; INTERNEGFC.
CC ProDom; PD002355; IRF. 1.
CC SMART; SM00348; IRF. 1.
CC PROSITE; PS00601; IRF. 1.
CC Transcription regulation; DNA-binding; Nuclear protein.
FT DNA BIND 16 118 TRYPTOPHAN PENTAD REPEAT.
FT DOMAIN 9 12 POLY-PRO.
FT DOMAIN 140 148 POLY-GLU.
SQ SEQUENCE 497 AA; 56005 MW; D8BD54DB946E354F CRC64;

Query Match 16.3%; Score 444.5; DB 1; Length 497;
Best Local Similarity 29.2%; Pred. No. 8.7e-19;
Matches 148; Conservative 58; Mismatches 188; Indels 113; Gaps 21;

QY 11 RVLFGWLGELTSSGCGYGLWLDDEARTCFRVPWKHFARKDLS-EADARIFKANAVAROR 69
DB 11 RVLFGWLGELTSSGCGYGLWLDDEARTCFRVPWKHFARKDLS-EADARIFKANAVAROR 69
QY 14 RVLKLPWLVQVNSCCYGLQWVNGEKKLFYIPWRHATRHGSPQDGDNTIFKAWAKETGK 73
DB 14 RVLKLPWLVQVNSCCYGLQWVNGEKKLFYIPWRHATRHGSPQDGDNTIFKAWAKETGK 73
QY 70 WPPSSRGCGPPEAETAEAGWKTNRFCALRSTRFRVWMLRDSGD-PADPHKVVYALSREL 128
DB 70 WPPSSRGCGPPEAETAEAGWKTNRFCALRSTRFRVWMLRDSGD-PADPHKVVYALSREL 128
QY 74 YTEG-----VDEADPAKWKANURCALNRSRDFQFYDGRDMPPOPKYKIY-----EV 120
DB 74 YTEG-----VDEADPAKWKANURCALNRSRDFQFYDGRDMPPOPKYKIY-----EV 120
QY 129 CWREGP-----GTDQTEAEAPAVPPPGPGPGPFLAHTAGLQAPGLPAP 175
DB 121 C-SNGPAPTESQTDYDVLGEEEEEEL-----QRMLPG--LSITEPAL--PGPPNAP 170
QY 176 AGDKGDLQLAQVQSCDLADHLITASWGADVPVTPKAFCEGQGLPLTGCAGGFGPLPAGEL 235
DB 176 AGDKGDLQLAQVQSCDLADHLITASWGADVPVTPKAFCEGQGLPLTGCAGGFGPLPAGEL 235
QY 171 -----YSLPKEDTKWPPALQPPVGL-----GPPVDPDNL 199
DB 171 -----YSLPKEDTKWPPALQPPVGL-----GPPVDPDNL 199
QY 236 YGWAVETTSPPGQ-----PAALTGGEAAPSPHOAEPLY-----SFPSPACTAVQEP 284
DB 236 YGWAVETTSPPGQ-----PAALTGGEAAPSPHOAEPLY-----SFPSPACTAVQEP 284
QY 200 L-----APPSGNPAGFRQLLPVLEFGPLASSQPP--TEPLLPDLLISPHMLPLT----- 247
DB 200 L-----APPSGNPAGFRQLLPVLEFGPLASSQPP--TEPLLPDLLISPHMLPLT----- 247

QY 285 SPGALDVTIMYKGRVTLQKVVGHP-SCTFLYGPDPAPVRAATDPO-----OVAPPS 333
DB 248 ---DLEIKFPQYRGRAPRTLTISNPGQCRLFYS-----QLEATQEQVELFGPVTLEQVRPS 300
QY 334 PAELPDQKQLRYTBEHLRHFVAPGLHLELRGQPLWARRMGCKVYWEVGGPPGSGASPTPA 393
DB 301 PEDIPSKQREYTWQLLDVLDRLGLILQOGDLYAIRLCQCKVFW--SGPCALAHGSCPN 358
QY 394 CLLPRNCPTPIPDFRVFQELVEFRARORRSPRYTYILYFGQDLSAGRPKESILVLYKL 453
DB 359 -PIQREVKTKLFSLEQFLNELILFQKQGTNTPTPPPEIFFCFGEWPPDVKPREKKLITQV 417
QY 454 EPWLCRVHLEGTQREGVSSLDSDLDL 480
DB 418 VFAARLLLEMFSGELSSWADSIRLQI 444

RESULT 7
IRF3_MOUSE
ID IRF3_MOUSE STANDARD; PRT; 419 AA.
AC P70671;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interferon regulatory factor 3 (IRF-3).
GN IRF3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Hakem R., Grossman A., Antonio L., Suggs S., Mak T.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH
CC OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT
CC FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT
CC (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1 (BY SIMILARITY).
CC 2- SUBCELLULAR LOCATION: Nuclear.
CC 3- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U75839; AAB36924.1; -
CC EMBL; U75840; AAB36925.1; -
CC HSSP; P23906; 2IRF.
CC MGD; MGI:1859179; IRF3.
CC InterPro; IPR001346; IRF.
CC Pfam; PF00605; IRF. 1.
CC PRINTS; PR00267; INTERNEGFC.
CC ProDom; PD002355; IRF. 1.
CC SMART; SM00348; IRF. 1.
CC PROSITE; PS00601; IRF. 1.
CC Transcription regulation; DNA-binding; Nuclear protein; Activator.
FT DNA BIND 7 107 TRYPTOPHAN PENTAD REPEAT.
SQ SEQUENCE 419 AA; 46852 MW; 1FF67C4E0FC7F027 CRC64;

Query Match 16.0%; Score 438; DB 1; Length 419;
Best Local Similarity 27.6%; Pred. No. 1.7e-18;
Matches 143; Conservative 56; Mismatches 182; Indels 138; Gaps 17;

QY 6 BRAPRVPLFGEVLLGEISGCGYGLWLDDEARTCFRVPWKHFARKDLS-EADARIFKANAV 65
DB 2 ETPKPRIL--PMLVQLDLQGLEGVANLDESTRFRIPWKHGLRQDAQWADFGIFQWAE 59
QY 66 ARGRWPPSSRGCGPPEAETAEAGWKTNRFCALRSTRFRVWMLRDSGDPAADPHKVVALS 125
```

Db 60 ASGAYT-----GKDKPDVST-----WKRNFRSALNRKEVLRLAADNSKDPYDHPKHYEFV 110
QY 126 RELCWREGPGT-----DQTEAEAPAAVPPGPGPGPLAHTAGLQAPGGLPAPAGD 178
Db 111 T-----PGARDPVHUGASPDNGKSSLSHSENLKLP-----DGLILGPLKD 153
QY 179 KGDLLQAOVQSCADLHLLTASWGADPVPTKAPGEGGLPLTACACAGGGLPAGELYGW 238
Db 154 EGSSDLAIV-----SDP----- 165
QY 239 AVETTPSPGPQPAALTTGEAAAPESPHQAEPYLSPPSACTAVQESPGL-----DV 291
Db 166 -SQQLSP-----NANFLNPAQ-----ENPLKQLLABEQWFEV 200
QY 292 TIMYKRTVLQKVGHPSCTFLXGPPDPVAVRATDQOVAFPSA-AELPDQKQLRYTEEL 350
Db 201 TAYRGRQVQQTFLCPGGLRVG--STADMTLPWQVTLDPDEGFLTDKLVEYGVQVL 258
QY 351 RHVAPGLHLRLGPFQWARMGKCKVYWEVG-----GPPGSASPTPACLLPRNCD 401
Db 259 KGLGNGLALWAGQCLWAQRLGHSHAFWALGEELLPDGSRGPDGE-----VHKDKD 309
QY 402 TPIDFRVFQELVEFRARORRSPRYTIVLGFQDLSAGRPKEKSLVVLKLEPWLVRVH 461
Db 310 GAVFDLRFPVADLTAFFMEGSGH-SPRYTLWFCMGMWPDQPVWVXRLVMVVKVVTCLKEL 368
QY 462 LEGTOREGVSSDLDLCLSSANSI-YDIECFELMEL 499
Db 369 LEMAREGASSLKTVDLHNSQFISLTSDQYKAYLQDL 407

RESULT 8

IRF6 MOUSE
ID IRF6 MOUSE STANDARD; PRT; 467 AA.
AC P97431;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 6 (IRF-6).
GN IRF6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Colon;
RA Groseman A., Mitterncker H.W., Antonio L., Mak T.W.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U73029; AAB36714.1; -
DR HSSP; P23906; 2IRF.
DR MGD; MGI:1859211; Irff6.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTFRNREGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 9 111 TRITOPHAN PENTAD REPEAT.
SQ SEQUENCE 467 AA; 53106 MW; 68CCAA90680FEDC8 CRC64;

Query Match 16.0%; Score 436; DB 1; Length 467;
Best Local Similarity 27.7%; Pred. No. 2.5e-18;
Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 17;
QY 1 MALAPRAAPRVLFGLWLGELISSCYEGQLWIDEARTCFRVPWKHFAR-KDISEADARI 59
Db 1 MALHPR-----RVLKLPWLVAQVDSGLYPLGLIWLHRDSKRFQIPWKHATRHSPQEEBENTI 56
QY 60 FKAWAVARGWRPSSSGGPPPEAEATAERAGWKTNERCALRSTRFRVLMRDNSGD-PADP 118
Db 57 FKAWAVETGKY-----QEGVDDDDPAK-----WKAQLRCALNKSRREFNLMDYDGTKVEPMNP 107
QY 119 HKVYALSRLCWREGPGTQTEAEAPAAVPPGPGPLAHTAGLQAPGGLPAPAGD 178
Db 108 VKIY-----QVC-----DIFQPG-----SVINPGSTGSAPWD 135
QY 179 KGDLLQAOVQSCADLHLLTASWGADPVPTKAPGEGGLPL-----TGACAGGGLPAGE 234
Db 136 EKNDVDEDEE-----EDELEQSQHHVPIQDTFPFLNINGSMPAPASVNGCNGCSP--- 188
QY 235 LYGNWVETTPSPGPQPAALTTGEAAAPESPHQAEPYLSPPSACTAVQESPGLDVTIM 294
Db 189 -----ESVMPKTEPLEMEVPOAPIQFFYSSPELWISSLPMT-----DLDIKFQ 231
QY 295 YKGRTVLQKV-VGHP-SCTFLYGPDP-----AVRATDQOVAFPSAPLDPQKQLRYT 346
Db 232 YRKEVGTQMTVSNPOGCLFYGDLPMPDQBELFGVSLQEVKFGPPEHITNEKQKLF 291
QY 347 EELLRHVAPGLHLRLGPFQWARMGKCKVYWEVGPPGSASPTPACLLPRNCDTPIFD 406
Db 292 SKLLDVMDRGLILEVSGHAIYAIRLCQCKVYWSGPCAPSLAAN-----LIERQKVKLFC 347
QY 407 FRVFFOELVEFRARORRSPRYTIVLGFQDLSAGRPKEKSLVVLKLEPWLVRVHLEGTQ 466
Db 348 LETFLSELIAHKGQIEKQPPFEIYLCFGEEDGPKLERKLLVQVPIPVVARNIEMFMS 407
QY 467 REGVSSDLDLCLSSANSIYDDIECFELMEL 499
Db 408 GDFTRSFDSGVRLQISTPD-INKNIVAQLKQL 439
RESULT 9
IRF6 HUMAN
ID IRF6 HUMAN STANDARD; PRT; 467 AA.
AC Q14856;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 6 (IRF-6).
GN IRF6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Groseman A., Mitterncker H.W., Antonio L., Ozato K., Mak T.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

CC the European Bioinformatics Institute. There are no restrictions on its
CG use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF027292; AAB84111.1; -;
DR EMBL; AL022398; CAA18545.1; -;
DR EMBL; BC014852; AAH14852.1; -;
DR HSSP; P23906; 2IRF.
DR Genew; HGNC:6121; IRF6.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD00267; INTERNEGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 9 111 TRYPTOPHAN PENTAD REPEAT.
SQ SEQUENCE 467 AA; 53129 MW; 7E28F5E0F5BA4053 CRC64;

Query Match 15.8%; Score 431; DB 1; Length 467;
Best Local Similarity 27.7%; Pred. No. 4.8e-18;
Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 20;

QY 1 MALAPERAAPRVLFGEWLLGEISSGCGEGLQWLDEARTCFRVPWKHFAR-KDLSEADARI 59
DB 1 MALHPR---RVRLKPLWVAQVDSGLYGLILWLRDSKRFQIPWKHATRHSPQEEENI 56

QY 60 FKAWAVARGWRPSSRGCGPPPEAEAEAGAKWTKNFCALSTRFRVWLRDMSGD-PADP 118
DB 57 FKAWAVETGKY---QEGVDDPPAK-----WKAQLRCALNKSREFNLMDYGTKEVPMNP 107

QY 119 HKVYALSRLCWREGCGTDTQTEAEAPVPPGCGPPGPFPLAHTAGLQAPGLPAPAGD 178
DB 108 VKIY---QVC-----DIPQFG-----SIINPGSTGAPWD 135

QY 179 KGDLLQAVQOQSLADHLHTASGADVPVTKAPGEGQGLPL-----TGACAGGPGLPAGE 234
DB 136 EXD---NDVDEDEDEDE-LDQSHVPIQDTPFLINSGSPAPASVGCNSVGCNSPEAV 191

QY 235 LYGWAVETPSPQPAALTTGEEAAPSPHQAEPYLSPSNACTAVQPSGALDVTIM 294
DB 192 ---W-----PKTEPLEMEVPO--APIQPFYSSPELWISSLFMT-----DLDIRQ 231

QY 295 YKGRTVLQKV-VGHP-SCTFLYGPPDP-----AVRATDPOQVAPSPAPLDPKOLRYT 346
DB 232 YKKEYGQVTWNSNPOGCRLLFYGDLPMPDQBELFGVSLQVKFGPEHITNEKQKLF 291

QY 347 EELLRHVAPGLHLELPGQPMARRMGKCKYVVEVGGPPGSGSPSTPACLLPRNCDTPIFD 406
DB 292 SKLLDVMDRGLILEVSGHAIYALRLCCKVYNSGCPACSLVAPN-----LIERQKKVKLFC 347

QY 407 FRVFPOLVEFRARQRGSPRYTIYLGQODLSAGRPKESLVLVKLEPWLRCRVHLEGTO 466
DB 348 LETFLSLDLIAHQKQIEKQKOPPPFIYLCFGEWPDGKPLERKILVQVIVPVARMIYEMS 407

QY 467 REGVSLDSDLDLCLSSANSLSVDDIECFMEL 499
DB 408 GDFTRFSDSGSVRLQISTPD-IKDNIVAKQL 439

RESULT 10
IRF4_MOUSE
ID IRF4_MOUSE STANDARD; PRT; 450 AA.
AC Q64287; Q60802;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 4 (IRF-4) (Lymphocyte specific interferon
DE regulatory factor) (LSIRF) (NF-EM5) (PU.1 interaction partner)
DE (Transcriptional activator PIP).
GN IRF4 OR SPIP.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/C;
RX MEDLINE=95317607; PubMed=7797077;
RA Eisenbeis C.F., Singh H., Storb U.;
RT "Pip, a novel IRF family member, is a lymphoid-specific,
RT PU.1-dependent transcriptional activator.";
RL Genes Dev. 9:1377-1387(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6, and 129/SvJ; TISSUE=Spleen;
RX MEDLINE=95334364; PubMed=7541907;
RA Matsuyama T., Grossman A., Mittrucker H.-W., Siderovski D.P.,
RA Kiefer F., Kawakami T., Richardson C.D., Taniguchi T., Yoshinaga S.K.,
RA Mak T.W.;
RT "Molecular cloning of LSIRF, a lymphoid-specific member of the
RT interferon regulatory factor family that binds the interferon-
RT stimulated response element (ISRE).";
RL Nucleic Acids Res. 23:2127-2136(1995).
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-
CC BINDS THE IMMUNOGLOBULIN LAMBDA LIGHT CHAIN ENHANCER, TOGETHER
CC WITH PU.1. PROBABLY PLAYS A ROLE IN ISRE-TARGETED SIGNAL
CC TRANSDUCTION MECHANISMS SPECIFIC TO LYMPHOID CELLS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER IN THE PRESENCE
CC OR ABSENCE OF A GLUTAMIC ACID RESIDUE.
CC -!- TISSUE SPECIFICITY: LYMPHOID CELLS.
CC -!- INDUCTION: NOT INDUCED BY INTERFERONS.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U34307; AAA75283.1; -;
DR EMBL; U11692; AAA75309.1; -;
DR EMBL; U20949; AAA75316.1; -;
DR EMBL; U20949; AAA75317.1; -;
DR HSSP; P23906; 2IRF.
DR MGD; MGI:1096873; Irfa.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERNEGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator;
KW Alternative splicing.
FT DNA_BIND 23 125 TRYPTOPHAN PENTAD REPEAT.
FT VARSPPLIC 165 165 MISSING (IN ISOFORM 2).
SQ SEQUENCE 450 AA; 51577 MW; 5FD94CA6C453869C CRC64;

Query Match 13.8%; Score 376; DB 1; Length 450;
Best Local Similarity 28.1%; Pred. No. 6.3e-15;
Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

QY 16 EWLGEISSGCGEGLQWLDEARTCFRVPWKHFARKDLS-EADARIFKAVAVARGRPSS 74
DB 26 QWLIDQIDSGKYPGLVWNEEKSVFRIPWKHAGKQDYNEEDAAALFKAWLFKGF--- 81

QY 75 RGGGPPPEAEAEAGAKWTKNFCALSTRFRVWLRDMSG-DPADPHKVVALSRELWCWREG 133
DB 82 REGIDKDPDPT-----WKTRLRCALNKSNDFELVRSQIDSDPYKYRI----- 127

QY 134 PGTDQTEAEAPAAVPPQGGPPG---FLATHAGLQAPGLPAPAGDKGDLILLQAVQOS 190
Db 128 -----VPEGAKGAKQLTLDQTQWAGHPYPMATP---YGSLLPAQVHNY 169
QY 191 CLADHLTASGADPVTAKPGQEGELPITGACAG---GGLP-----AGELYGWAVE 241
Db 170 MPPH---DRSW-RDYAPDQSHPEIPYQCPVTFGRHHWQSPSCENGCGQVTGTGYACAPP 226
QY 242 TTPSPG-PQPAALTGGAAPESPQAEPLSPSPSACTAVQSPSGALDVTIMYKGRV 300
Db 227 ESQAPGPIPSIRSABALA-----LSDC-----RLHICLYYRDILV 263
QY 301 LQKVVGHP-SCTFLYGPDPDAVRATDPOQVAFPSPAELPDQKOLRYTEELLRHVAPGLHL 359
Db 264 KELTTTSPGCRISHG---HTYDVSNLDQVLFYP---DNGQRKNIKELLSHLERGLVL 317
QY 360 ELRGQPLWARMGKCKYVWVGPPGSGAS---PSTPACLLPRNCPTPIFDPRVFQELVEF 417
Db 318 WMAPDGLYAKRLCOSRIYD---GPLALCSDRPNK---LERDQCKLFDQQLSELOVF 371
QY 418 RARQRGSPRYTYLGFQDLSAGRPKESLVLKLEPWLCR 459
Db 372 -AHGRPAPRFQVTLCEGEEFPDPQ-RQRKLITAHVEPLLAR 411

RESULT 11
IRF4 HUMAN
ID IRF4 HUMAN STANDARD; PRT; 451 AA.
AC Q15306; Q99660;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 4 (IRF-4) (Lymphocyte specific interferon
regulatory factor) (LSIRF) (NF-EM5) (Multiple myeloma oncogene 1).
GN IRF4 OR MUM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Spleen;
RX MEDLINE=97079690; PubMed=8921401;
RA Grossman A., Mittrucker H.W., Nicholl J., Suzuki A., Chung S.,
RA Antonio L., Sugga S., Sutherland G.R., Siderovski D.P., Mak T.W.;
RT "Cloning of human lymphocyte-specific interferon regulatory factor
4 (LSIRF/hIRF4) and mapping of the gene to 6p23-p25.";
RL Genomics 37:229-233(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Spleen;
RA Iida S., Butler M., Hatzivassiliou G., Pulivarthi R., Klein B.,
RA Corradini P., Boccadoro M., Nilsson K., Chaganti R.S.K.,
RA Dalla-Favera R.;
RT "Deregulated expression of MUM1 gene by t(6;14) (p25;q32) chromosomal
translocation in multiple myeloma.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-
STIMULATED RESPONSE ELEMENT (ISRE) OF THE MHC CLASS I PROMOTER.
CC BINDS THE IMMUNOGLOBULIN LAMBDA LIGHT CHAIN ENHANCER, TOGETHER
CC WITH PU.1. PROBABLY PLAYS A ROLE IN ISRE-TARGETED SIGNAL
CC TRANSDUCTION MECHANISMS SPECIFIC TO LYMPHOID CELLS.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER IN THE PRESENCE
CC OR ABSENCE OF A GLUTAMIC ACID RESIDUE.
CC -I- TISSUE SPECIFICITY: LYMPHOID CELLS.

CC -I- INDUCTION: NOT INDUCED BY INTERFERONS.
CC -I- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U52682; AAC50779.1; -;
DR EMBL; U63738; AAB37258.1; -;
DR EMBL; BC015752; AAH15752.1; -;
DR HSSP; P23906; 2IRF.
DR TRANSFAC; T04929; -;
DR Genew; HGNC:6119; IRF4.
DR MIM; 601900; -;
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.1.
DR PRINTS; PR00267; INTERREGFCT.
DR ProDom; PD002355; IRF.1.
DR SMART; SM00348; IRF.1.
DR PROSITE; PS00601; IRF.1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator;
KW Alternative splicing.
FT DNA_BIND 23 125 TRYPTOPHAN PENTAD REPEAT.
FT VARSPPLIC 165 165 MISSING (IN ISOFORM 2).
FT CONFLICT 300 300 Q -> H (IN REF. 2).
FT CONFLICT 306 306 K -> N (IN REF. 2).
FT CONFLICT 333 333 R -> T (IN REF. 2).
SQ SEQUENCE 451 AA; 51772 MW; 17CD1327C6F5BF6A CRC64;
Query Match 13.3%; Score 363; DB 1; Length 451;
Best Local Similarity 27.5%; Pred. No. 3.5e-14;
Matches 136; Conservative 73; Mismatches 193; Indels 92; Gaps 24;
QY 16 EWLGLISSGCVGEGLOWLDEARTCFRVPWKHFARKDLS-EADARIFKAWAVGRWPPSS 74
Db 26 QWLIDQIDSGKVPGLVWENEKSIIFRIPWKHAGKQDYNEEDAAALFKAWLFGKGF--- 81
QY 75 RGGGPPPEAETAERACWKTNFRCALSTRFRVMLRDNNG-DPADPHKVYALSRELWREG 133
Db 82 REGIDKPDPT-----WKRLRCALNKSNDFELVRSQDIDSDPKYVRIVPEGAK-- 134
QY 134 PGTDQTEAEAPAAVPPQGGPPGFLATHAGLQAPGP-LPAPAGDKGDLILLQAVQOSCL 192
Db 135 -GAKQLTLEDPPQ-----SMSPYPT-MTTPYPSLPA-----OOVHNYMM 171
QY 193 ADHLTASGADPVTAKPGQEGELPIT-----GACAGGPGLPAGELYGWAVET 242
Db 172 PP--LDRSW-RDYVDPQHPETPYQCPMTFGRGHWHQGPACENGQV-TGTFYACAPPE 227
QY 243 TTPSPG-PQPAALTGGAAPESPQAEPLSPSPSACTAVQSPSGALDVTIMYKGRVTL 301
Db 228 SQAPGVPTPEPSIRSABALA-----FSDC-----RLHICLYYRETLVK 264
QY 302 QKVVGHP-SCTFLYGPDPDAVRATDPOQVAFPSPAELPDQKOLRYTEELLRHVAPGLHLE 360
Db 265 ELTTSPEGCRISHG---HTYDVSNLDQVLFYP---EDNGQRKNIKELLSHLERGVVLW 318
QY 361 LRGQPLWARMGKCKYVWVGPPGSGAS---PSTPACLLPRNCPTPIFDPRVFQELVEPR 418
Db 319 MAPDGLYAKRLCOSRIYD---GPLALCSDRPNK---LERDQCKLFDQQLSELOAF- 371
QY 419 ARQRGSPRYTYLGFQDLSAGRPKESLVLKLEPWLRCRVHLEGTQREGVSSLDSSDL 478
Db 372 AHGRSLPRFQVTLCEGEEFPDPQ-RQRKLITAHVEPLLARQLYFYFAQQNSGHFLRGYDL 430
QY 479 DLCLSSANSLYDDI 492
Db 431 PEHISNPEDYHRSI 444

QY	304	VGHPSCTFLYG-----PPDAVATDP-----QVAFPSAELPDOKQLRYTEELLRHV	353
DB	217	LVGHITTSYPEGCRSLSQSPNHEKLYTPDSLEHVFPSAEAIQNDROKQITKCLFGHL	276
QY	354	APGLHLELRGQLWARMGKCKVYVEVGGPPGSPSTPACLLPRNCDDTPIFDFRVFFOE	413
DB	277	ERGVLHNSKOGIFIKRLCQGRVFWGNTVYKDRPSK-----LDRDEVKIFDTNLFRE	332
QY	414	LVEFRARQGRSPRYTYLFGQDLSAGRPKEKSLVLVLEPMLCRLVHLEGTQREGVSSL	473
DB	333	LQQYNNQGR-FPDSRVMLCGEFPDTPVLRCKLILVQEQLCVQVMEAGKTCSSPM	391
QY	474	DSSDL-----DLC	481
DB	392	LPDDVQEQVYRIFQDIC	409
<p>RESULT 13</p> <p>ICSB MOUSE</p> <p>ID ICSB MOUSE STANDARD; PRT: 424 AA.</p> <p>AC P23611;</p> <p>DT 01-NOV-1991 (Rel. 20, Created)</p> <p>DT 01-NOV-1997 (Rel. 35, Last sequence update)</p> <p>DT 01-NOV-1997 (Rel. 35, Last annotation update)</p> <p>DE Interferon consensus sequence binding protein (ICSBP).</p> <p>GN ICSBP1 OR ICSBP.</p> <p>OS Mus musculus (Mouse).</p> <p>OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;</p> <p>OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.</p> <p>OX NCBI_TaxID=10090;</p> <p>RN [1]</p> <p>RP SEQUENCE FROM N.A.</p> <p>RX MEDLINE=90251633; PubMed=2111015;</p> <p>RA Driggers P.H., Ennist D.L., Gleason S.L., Mak W.-H., Marks M.S.,</p> <p>RA Levi B.-Z., Flanagan J.R., Appella E., Ozato K.;</p> <p>RT "An interferon gamma-regulated protein that binds the interferon-</p> <p>RT inducible enhancer element of major histocompatibility complex class</p> <p>RT I genes.";</p> <p>RL Proc. Natl. Acad. Sci. U.S.A. 87:3743-3747(1990).</p> <p>CC -!- FUNCTION: SPECIFICALLY BINDS TO THE UPSTREAM REGULATORY REGION</p> <p>CC OF TYPE I IFN AND IFN-INDUCIBLE MHC CLASS I GENES (THE INTERFERON</p> <p>CC CONSENSUS SEQUENCE (ICS)). PLAYS A REGULATORY ROLE IN CELLS OF THE</p> <p>CC IMMUNE SYSTEM.</p> <p>CC -!- SUBCELLULAR LOCATION: Nuclear.</p> <p>CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN LYMPHOID TISSUES.</p> <p>CC -!- INDUCTION: BY INTERFERON GAMMA.</p> <p>CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.</p> <p>CC -----</p> <p>CC This SWISS-PROT entry is copyright. It is produced through a collaboration</p> <p>CC between the Swiss Institute of Bioinformatics and the EMBL outstation</p> <p>CC the European Bioinformatics Institute. There are no restrictions on its</p> <p>CC use by non-profit institutions as long as its content is in no way</p> <p>CC modified and this statement is not removed. Usage by and for commercial</p> <p>CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)</p> <p>CC or send an email to license@isb-sib.ch</p> <p>CC -----</p> <p>DR EMBL; M32489; AAA37878.1; .</p> <p>DR PIR; A35861; A35861.</p> <p>DR HSSP; P23906; 2IRF.</p> <p>DR TRANSFAC; T00402; .</p> <p>DR MGD; MGI:96395; Icbdp.</p> <p>DR InterPro; IPR001346; IRF.</p> <p>DR Pfam; PF00605; IRF; 1.</p> <p>DR PRINTS; PR00267; INTERREGFCT.</p> <p>DR ProDom; PD002355; IRF; 1.</p> <p>DR SMART; SM00348; IRF; 1.</p> <p>DR PROSITE; PS00601; IRF; 1.</p> <p>DR Transcription regulation; DNA-binding; Activator; Nuclear protein;</p> <p>DR Interferon induction.</p> <p>DR DNA BIND 9 110 TRYPTOPHAN PENTAD REPEAT.</p> <p>CC SEQUENCE 424 AA; 48237 MW; FBE79A768468EB2 CRC64;</p>			

Query Match	12.2%;	Score 333.5;	DB 1;	Length 424;
Best Local Similarity	24.3%;	Pred. No. 1.6e-12;		
Matches	121;	Conservative	65;	Mismatches 152; Indels 159; Gaps 19;
QY	16	EWLLGRISSCYEGLWLDDEARTCFRVPWKHFARKULS-BADARIFKAWAVARGRPPSS	74	
Db	12	QWLEIQIDSSMYGLIWEDEMTFRIPWKHAGQYDQNEVDASIFKAWAVFKGF----	67	
QY	75	RGGGPPEAETARAGKTNFRALSTRFRVLMRNSG-DPADPHKVALSRELWCWREG	133	
Db	68	-----KEGDKAPATWKTRLCALNKSPDFEETDRSQDIDSEPKYVRI-----	112	
QY	134	PGTDQTEAEAPAAVPPPPQGGPFFLAHTHAGLQAPGLPAPAGDKGDLILLQAVQOSCLA	193	
Db	113	--VPDEEKCKLGV-----APAG-----CMS	131	
QY	194	DHLLTASWAGDPVPTKAPGEGQ-EGILPTGACAGGPGLPAGELYGWAVETTPSPGPQPA	252	
Db	132	E-----VPDEMGGRSEIEELIKE-----PSVDEYMGWTKRSPSP-----	165	
QY	253	LTTGERAAAPSPHQAEPYLSFSPSACTA-----VOERSPG-----ALD-----	290	
Db	166	-----PEACKSQILPDWVQQPSAGLPLVTGYAAYTHHSAPS	203	
QY	291	---VTIMYKGRITLQKV-----VGHPSCTFLYPPDPDAVRATDPOQVAFPPSPAE	337	
Db	204	QMWISFYGGKLVQATTTCLGCRULSQPLKLYGPD-----GLEP--VCFPTADTI	256	
QY	338	PDQKQRYTEBLRHVAPLHLRGLWARMGKVKVWVGPPGSPASPTPACLLP	397	
Db	257	PSERQOVTRKLFCHLERGVLLHSNRKGVFVKRLCQGRVFCSGNAVVCGRPNK----	312	
QY	398	RNCDTPIEDFRVFQELVEFRARQGRSPRYTIVLFGQDLSAGRPKESLVLKLEPWL	457	
Db	313	RDEVQVDFNQPIRELQQYATQSR-LPDSRVVLCFGEETVPLRSKLILVQVEQLY	371	
QY	458	CRVHLEGTQRE-GVSSL 473		
Db	372	ARQVLEEAGKSCGAGSL 388		
RESULT 14				
IRTF HUMAN				
ID	IRTF HUMAN	STANDARD;	PRT;	393 AA.
AC	Q00978;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Transcriptional regulator ISGF3 gamma subunit (IFN-alpha responsive			
DE	transcription factor subunit) (interferon stimulated gene factor 3			
DE	gamma) (ISGF3 p48 subunit) (ISGF-3 gamma).			
GN	ISGF3G.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=92334329; PubMed=1630447;			
RA	Veals S.A.; Schindler C.; Leonard D.G.B.; Fu X.-Y.; Aebersold R.H.,			
RA	Darnell J.E. Jr.; Levy D.E.;			
RT	"Subunit of an alpha-interferon-responsive transcription factor is			
RT	related to interferon regulatory factor and myb families of			
RT	DNA-binding proteins.";			
RL	Mol. Cell. Biol. 12:3315-3324(1992).			
CC	-1- FUNCTION: RESPONSIBLE FOR THE INITIAL STIMULATION OF INF-ALPHA-			
CC	RESPONSIVE GENES. IT RECOGNIZES AND BINDS TO THE INF-STIMULATED			
CC	RESPONSE ELEMENT, OR ISRE WITHIN THE REGULATORY SEQUENCES OF			
CC	TARGET GENES. ISGF3 PLAYS A PRIMARY ROLE IN THE TRANSMISSION OF A			
CC	SIGNAL FROM THE CELL SURFACE TO THE NUCLEUS.			
CC	-1- SUBUNIT: ACTIVE ISGF3 CONSISTS OF TWO COMPONENTS, A REGULATORY			
CC	MULTISUBUNIT COMPLEX, ISGF3 ALPHA (COMPOSED OF STAT1 ALPHA/BETA,			

CC	AND STAT2, ALSO DESIGNATED P91/P84 AND P113 RESPECTIVELY) AND			
CC	ISGF3 GAMMA (OR P48). THEY ASSEMBLE WHEN THE ISGF3 ALPHA COMPONENT			
CC	IS PHOSPHORYLATED VIA INF ACTIVATION.			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- INDUCTION: BY IFN-ALPHA AND IFN-BETA. UPON STIMULATION THE			
CC	REGULATORY PHOSPHORYLATED ALPHA AND BETA SUBUNITS ASSEMBLE WITH			
CC	THE GAMMA SUBUNIT AND TRANSLOCATE FROM THE CYTOPLASM TO THE			
CC	NUCLEUS.			
CC	-1- SIMILARITY: BELONGS TO THE IRF FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M87503; AAA58687.1; -.			
DR	PIR; A45017; A45017.			
DR	HSSP; P23906; 2IRF.			
DR	TRANSFAC; T01456; -.			
DR	Genew; HGNC:6131; ISGF3G.			
DR	MIM; 147574; -.			
DR	InterPro; IPR001346; IRF.			
DR	Pfam; PF00605; IRF; 1.			
DR	PRINTS; PR00267; INTERREGFCT.			
DR	ProDom; PD002355; IRF; 1.			
DR	SMART; SM00348; IRF; 1.			
DR	PROSITE; PS00601; IRF; 1.			
KW	Transcription regulation; DNA-binding; Activator; Nuclear protein;			
KW	Interferon induction.			
FT	DOMAIN 11 112 TRYPTOPHAN PENTAD REPEAT.			
FT	DOMAIN 181 187 POLY-SER.			
SQ	SEQUENCE 393 AA; 43696 MW; F8E3784354BFD4A9 CRC64;			
Query Match 12.0%; Score 328.5; DB 1; Length 393;				
Best Local Similarity 25.7%; Pred. No. 2.8e-12;				
Matches 120; Conservative 69; Mismatches 179; Indels 99; Gaps 19;				
QY	17	WLLGEISSCYEGLWLDDEARTCFRVPWKHFARKULS-ADARIFKAWAVARGRPPSSR	75	
Db	15	WVVEQVESGQFFGVCCWDDTAKTMFRIPWKHAGQDFREDODAAFFKAWAIFKGY-----	69	
QY	76	GGGPPPEAETARAGKTNFRALSTRFRVLMRNSG-DPADPHKVALSRELWCWREGP	134	
Db	70	-----KEGDTGGPAVWKTRLCALNKSSSEKVEPERGRMDVAEPKYQLLPFGIVSGQP	124	
QY	135	GTDDQTEAEAPAAVPPPPQGGPFGPFLAHTHAGLQAPGLPAPAGDKGDLILLQAVQOSCLAD	194	
Db	125	GTQKVPSK-----RQHSSVS-----SERKEED---AMQNCTLSP	156	
QY	195	HLTASWAGDPVPTKAPGEGQGLPITGACAGGPGLPAGELYGWAVETTPSPGPQPAALT	254	
Db	157	SVLQDSLNNEE-----EGASG-----GAVHSDIG-----SSSSSSSPPEQVETDT	196	
QY	255	TGEAAAPESPQAE-----PVLSPSPSACTAVQSPSPG-ALDVTIMYKGRITLQKVGHPS	309	
Db	197	T-----EAPFQDQSRSLFLLP-----PEPDSLLITFYNGRVVGAQVQSLD	240	
QY	310	CTFLYGPDPFAVRATDPOQVAFPPSPAEPLPDQKQRYTEBLRHVAPGLHLELPGQLMAR	369	
Db	241	CRIV--ABPSGSESSMEQVLPFKPGLP-----TORLLSQLERGILVASPRGLFVQ	291	
QY	370	RMGKCKVYEVGGPPGSPASPTPACLLPRNCDTPIEDFRVFQELVEFRARQRRG-SPRY	428	
Db	292	RLCPIPISNAPQAPFGPGFH-----LLPSNECELEFRTAYFCRDLV--RYFQGLGPPPKF	345	
QY	429	TIYVFGQDLSAGRPKESLVLKLEPWLCRVHLEGT--OREGVSSL 473		
Db	346	QVTLNFWESHSSHTPQNLIIVKMEQAFARYLLEQTPEQQAAILSL 392		

RESULT 15	ICSB_HUMAN	STANDARD;	PRT;	426 AA.
ID	ICSB_HUMAN			
AC	Q02556;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Interferon consensus sequence binding protein (ICSBP).			
GN	ICSBP1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
ON	NCBI_TaxID=9606;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung, and Monocytes;			
RX	MEDLINE=93094284; PubMed=1460054;			
RA	Weisz A., Marx P., Sharf R., Appella E., Driggers P.H., Ozato K.,			
RA	Levi B.-Z.;			
RT	"Human interferon consensus sequence binding protein is a negative			
RT	regulator of enhancer elements common to interferon-inducible			
RT	genes.";			
RL	J. Biol. Chem. 267:25589-25596(1992).			
RN	[2]			
RA	REVISIONS.			
RP	Schmidt M.;			
RL	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: SPECIFICALLY BINDS TO THE UPSTREAM REGULATORY REGION			
CC	OF TYPE I IFN AND IFN-INDUCIBLE MHC CLASS I GENES (THE INTERFERON			
CC	CONSENSUS SEQUENCE (ICS)). PLAYS A NEGATIVE REGULATORY ROLE IN			
CC	CELLS OF THE IMMUNE SYSTEM.			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- TISSUE SPECIFICITY: PREDOMINANTLY IN LYMPHOID TISSUES.			
CC	-1- INDUCTION: BY INTERFERON GAMMA.			
CC	-1- SIMILARITY: BELONGS TO THE IRF FAMILY.			
CC	CC			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or			
CC	send an email to license@isb-sib.ch).			
CC	EMBL; M91196; AAB63813.1; --			
DR	FIR; A45064; A45064.			
DR	HSP; P23906; 2IRF.			
DR	TRANSFAC; T02038; --			
DR	Genew; HGNC:5358; ICSBP1.			
DR	MIM; 601565; --			
DR	InterPro; IPR001346; IRF.			
DR	Fam; PF00605; IRF; 1.			
DR	PRINTS; PR00267; INTPRNREGFCT.			
DR	ProDom; PD002355; IRF; 1.			
DR	SMART; SM00348; IRF; 1.			
DR	PROSITE; PS00601; IRF; 1.			
KW	transcription regulation; DNA-binding; Nuclear protein;			
KW	interferon induction.			
FT	DNA BIND 9 110			
FT	SEQUENCE 426 AA; 48356 MW; 1535D1B7C83E0355 CRC64;			
Query Match	11.3%; Score 308.5; DB 1; Length 426;			
Best Local Similarity	21.8%; Pred. No. 4.2e-11;			
Matches 116;	Conservative 71; Mismatches 161; Indels 185; Gaps 18;			
QY	16 EWLLGISSGCGYBGLWLDERTCFRVPWKHFARKOLS-BADARIFKAWAVARGWPPSS 74			
Db	12 QWLIEQIDSSWYPLGIWENEKSNFRIPWKHAGKQDYNQVDSIFKAWAVFKGF ---- 67			
QY	75 RGGGPPPEAETABRAGKTNFRCALRSTRFRVMLRDNSG-DPADPHKVYALSRELCLWREG 133			
Db	68 -----KEGDAKPAWTKRLCALNKSPPFEVTDTSQIDISBPYKVYRI----- 112			

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:43:22 ; Search time 36 Seconds
(without alignments)
2878.936 Million cell updates/sec

Title: US-09-647-965-9
Perfect score: 2731
Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSLYDDIECFLEMELEQFA 503

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Description
1	1116	40.9	Q9ue79 homo sapien
2	778	28.5	Q90zd4 gallus gall
3	603	22.1	Q9hb64 homo sapien
4	440	16.1	Q97578 xenopus lae
5	430	15.7	Q9n136 ovis aries
6	428	15.7	Q91vd0 mus musculus
7	420	15.4	Q8wnq4 sus scrofa
8	364.5	13.3	Q96gl3 homo sapien
9	364	13.3	Q98tx7 gallus gall
10	362	13.3	Q98tx6 gallus gall
11	344	12.6	Q99419 homo sapien
12	321	11.8	Q90wi0 gallus gall
13	281.5	10.3	Q9qz17 mus musculus
14	194	7.1	Q924t6 mus musculus
15	183	6.7	Q9rkr9 streptomyce
16	177.5	6.5	Q98tal fugu rubrip

17	176	6.4	1146	13	Q90584
18	173.5	6.4	322	6	Q9EG75
19	169.5	6.2	349	6	Q9n137
20	168.5	6.2	297	13	Q9pwi9
21	166	6.1	349	11	Q923E9
22	166	6.1	349	11	Q8R4S0
23	165.5	6.1	469	12	Q92270
24	165	6.0	3179	12	Q8V2A4
25	164	6.0	315	13	Q9I838
26	163	6.0	91	6	Q95290
27	163	6.0	890	11	Q921A0
28	162.5	6.0	585	12	Q41935
29	162	5.9	98	13	Q90WWS
30	162	5.9	1660	5	Q9NKN0
31	161.5	5.9	551	4	Q9Y6I3
32	161	5.9	889	16	Q9F2N5
33	160.5	5.9	1400	5	Q9NA90
34	159	5.8	576	16	Q9XA04
35	158.5	5.8	694	6	Q9GKY7
36	158.5	5.8	839	10	Q9SN46
37	158.5	5.8	1463	16	Q9ADP6
38	157.5	5.8	3247	12	Q65553
39	156	5.7	381	5	Q94399
40	156	5.7	552	4	Q9BR39
41	155	5.7	576	4	Q9HA18
42	155	5.7	678	11	Q8QZT8
43	155	5.7	801	5	Q23635
44	154	5.6	696	11	Q9ET79
45	154	5.6	1691	11	Q9ESQ2

ALIGNMENTS

RESULT 1

Q9UE79 PRELIMINARY; PRT; 210 AA.
AC Q9UE79;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE Putative interferon regulatory factor 7C.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND BREAST;
RX MEDLINE=97459673; PubMed=9315633;
RA Zhang L., Pagano J.S.;
RT "IRF-7, a new interferon regulatory factor associated with Epstein-Barr virus latency";
RL Mol. Cell. Biol. 17:5748-5757(1997).
DR EMBL; U53832; AAB80691.1; --
SQ SEQUENCE 210 AA; 23592 MW; 45895671CEAA18F5 CRC64;

Query Match 40.9%; Score 1116; DB 4; Length 210;
Best Local Similarity 99.0%; Pred. No. 9.5e-70;
Matches 208; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	294	MYKGRVTQKVVGHPSCTFLYGPDPVAVRATPQQVAFPSPAELPQKQLRYTEELLRHV	353
Db	1	MYKGRVTQKVVGHPSCTFLYGPDPVAVRATPQQVAFPSPAELPQKQLRYTEELLRHV	60
QY	354	APGLHLELRGPQWARMGMCKVYWEVGGPPGSPSTPACLLPRNCDTPIFDFRVFFQE	413
Db	61	APGLHLELRGPQWARMGMCKVYWEVGGPPGSPSTPACLLPRNCDTPIFDFRVFFQE	120
QY	414	LVEFRARQGRGSPRTIYILFGQDLSAGRPKESLVLVKLEPWLCHRVHLEGTOREGVSSL	473
Db	121	LVEFRARQGRGSPRTIYILFGQDLSAGRPKESLVLVKLEPWLCHRVHLEGTOREGVSSL	180

```
QY 474 DSDLDLCLSSANSLYDDIECFLEMELEQPA 503
DB 181 DSSSLSLCLSSANSLYDDIECFLEMELEQPA 210

RESULT 2
Q90ZD4 PRELIMINARY; PRT; 491 AA.
AC Q90ZD4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Interferon regulatory factor-3.
GN IRF-3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA May D.L., Grant C.E., Deeley R.G.;
RT "Cloning and Promoter Analysis of the Chicken Interferon Regulatory
RT Factor-3 Gene.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268079; AAK58583.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.
DR ProDom; PD002355; IRF.
DR ProSITE; PS00601; IRF; 1.
DR SMART; SM00348; IRF; 1.
SQ SEQUENCE 491 AA; 54440 MW; 90E0C80A2624B1B4 CRC64;

Query Match 28.5%; Score 778; DB 13; Length 491;
Best Local Similarity 38.5%; Pred. No. 5.5e-46;
Matches 208; Conservative 59; Mismatches 181; Indels 92; Gaps 17;

QY 2 ALAPRAAPRVLFGEWLLGHSICGVEGLQWLDEARTCFRPVWKHFARKDLSADARIFKAWA 61
DB 3 ALDSEGAQKLRFGPMLNVASSLGRLGCLWIDPDRRIPIPKHNARKDVTSSDVEIFK 62
QY 62 AWAVAGRWPPSSRGSGGPPPEATATAGWKTFRFCALRSTRFPVWMLRDNSGDPADPHKV 121
DB 63 AWAKASGRY-----EGNAEDPAKWKTFRFCALRSTHFMFLLEDKSCNDPDPHKV 111
QY 122 YALSRELWCWEGPGTQDTQAEAPAAVPPPO.-GGP-----PGPFLAHTHAGLO--- 167
DB 112 YAVA-----SGVPNDRSGSPVAGALQQQPLLNNHDLAENTPT 152
QY 168 -AFGPLPAPAGDKDL-LLQAVQSCSLADHLLTASMGADVPVTKAPGEG--QEGLEPLTG- 222
DB 153 DSTEGVAAALATQVLDLQLQVLQHCNISAL-----GSQPTLWAHTGDALPEDALLPGQ 207
QY 223 -ACAGGGLPAGLYGW-AVETTPSGPOP-----AALTGEAAAPSPHQAEP 269
DB 208 DGLCPGP-----QFQWRQLQEEPLLGNQPLTGGCGQDGAGALPVSECAIPAPSPAE 262
QY 270 YL-----SPSPSACTAVQEPSFGALDVTIMYKGRVQLQKVGHPSCTFLYGPDPAPVRATD 325
DB 263 LLFQSANPAPPAGDIGGLPPLLDITYYRGKVVYQEQVDDSRCLVAYQPLDPAV--AE 320
QY 326 PQQVAFPSPAELPQKQLRYTEELLRHVAPGLHLELPGQLWARRMGKCKVYWEV----- 380
DB 321 QRLVLFPSPASLPDRQRRTYNTNLE--VAGLRLEQRAGQLLATRLKCKVFWALSQOLE 378
QY 381 GGPFGSASPTPACLLPRNCDTIFDPRVFQELVEFRAPQRGRSPRYTIYLFQGDLSA 440
DB 379 GGEP-----PLNLLHRDQETTFIDPRVFCTELRDRFRDSRRERSPDFTIFLCFGQCFSS 431
QY 441 GRPKESLVLVKLEPMLCRVHLEGTQREGVSSLDSDSLDLCSSANSLYDDIECFLEMELE 500
DB 432 TKPKESKLLVLKLPVQCFEYVQVQVQGGASSLNSGNVLSQLSDSNFLFELIEQYHMQTD 491
```

```
RESULT 3
Q9HB64 PRELIMINARY; PRT; 128 AA.
AC Q9HB64;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Interferon regulatory factor-7H (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20493557; PubMed=10924517;
RA Lu R., Au W.-C., Yeow W.-S., Hageman N., Pitha P.M.;
RT "Regulation of the promoter activity of interferon regulatory factor-7
RT gene. ACTIVATION BY INTERFERON AND SILENCING BY HYPERMETHYLATION.";
RL J. Biol. Chem. 275:31805-31812(2000).
DR EMBL; AF277159; AAG30003.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR ProSITE; PS00601; IRF; 1.
DR NON_TER 128 128
SQ SEQUENCE 128 AA; 14363 MW; 09B5B6F3CC11BB15 CRC64;

Query Match 22.1%; Score 603; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.5e-34;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RAAPRVLFGEWLLGHSICGVEGLQWLDEARTCFRPVWKHFARKDLSADARIFKAWA 66
DB 20 RAAPRVLFGEWLLGHSICGVEGLQWLDEARTCFRPVWKHFARKDLSADARIFKAWA 79
QY 67 RGRWPPSSRGSGGPPPEATATAGWKTFRFCALRSTRFPVWMLRDNSGDP 115
DB 80 RGRWPPSSRGSGGPPPEATATAGWKTFRFCALRSTRFPVWMLRDNSGDP 128

RESULT 4
Q57578 PRELIMINARY; PRT; 459 AA.
AC Q57578;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE XIRF-6 protein.
GN XIRF-6.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98086218; PubMed=9426249;
RA Hatada S., Kinoshita M., Takahashi S., Nishihara R., Sakumoto H.,
RA Fukui A., Noda M., Asashima M.;
RT "An interferon regulatory factor-related gene (XIRF-6) is expressed in
RT the posterior mesoderm during the early development of Xenopus
RT laevis.";
RL Gene 203:183-188(1997).
DR EMBL; D86492; BAA24349.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
```

```
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 459 AA; 52475 MW; 306DA77C6586871C CRC64;

Query Match 16.1%; Score 440; DB 13; Length 459;
Best Local Similarity 27.1%; Pred. No. 1.1e-22;
Matches 140; Conservative 74; Mismatches 197; Indels 106; Gaps 19;

QY 1 MALAPERAAPRVLFGEWLLGEISGCGYEGLOWLDEARTCFRVPWKHFAR-KDLSADARI 59
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 1 MAHPR-----RVRLKPWLVAQVDSGLYPLGLIWLHRDSKRFQIPWKHATRHSPQEEENTI 56
QY 60 FKAWAVAGRWPPSSRGGGPPPEAEATRAAGKTNFRCAIRSTRFVLMRDNSSG-PADP 118
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 57 FKAWAVETGKYQEG-----ADEPDAKWAQIRCALNKSREFKLMYDGTKEVPMNP 107
QY 119 HKVYALSRELWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHAGLQAPGLPAPAGD 178
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 108 VKIY-----EVC-----DIPQSG-----SIINPGSTGSPWD 135
QY 179 KGDLLQVQOCLADHLLTASWGADPVPTKAP-----GEGQGLPLTGACAGGGLPA 232
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 136 DDDFEADLNQ-----QNHVPISEPFNCLINDSPISSTGSGCT----- 176
QY 233 GELYGWAVETTPSGQPPAAL--TTGEAAAPSPHQAEPYLSPSACTAVQPSGALD 290
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 177 -----PEQTWPKTEPQEMEVPTSGPADFFSP-----EMWISLEMT-----DLE 217
QY 291 VTIMYKGRVLOKV-VGHP-SCTFLYGGPPDPVAVRATD-----PQVAPPSPAPLDPQKQ 342
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 218 IQFYRGKMGQTMVSNPQGCRLFGDLGPNQEELEFGPITLQVRFPGTEQIVNEKQ 277
QY 343 LRYTELLRHVAPGLHLELRGPMWARRMGKCKVYVEGPPGSGASPSPTPACLLPRNCDT 402
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 278 KLFTSKLLDVMRGILLEVSHAIYAIRLCQCKVYWGSCAPSLVAPN-----LIERQKKVK 344
QY 403 PFDFRVFQELVEFARQRGRSPRYTIYLGFGQDLSAGRPKESILVVKLEPWLRCVHL 462
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 334 KLFCVETFLSLDLSHQKGIITQKPPPEIYLGFGEEWPDGKPKERKLIIVQIIPVARMII 393
QY 463 EGTQREGVSSLDLCLSSANSYDDIECFMEL 499
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 394 EMFTGDSFSGSIRLQISIPD-IKDNIVSHLKH 429

RESULT 5
Q9N136 PRELIMINARY; PRT; 467 AA.
AC Q9N136
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Interferon regulatory factor 6.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=15 DAY PREGNANT UTERUS;
RA Choi Y., Spencer T.E., Bazer F.W.;
RT "Cloning and Analysis of Ovine IRF-6.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF228446; AAF34782.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERFERGFCF.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 467 AA; 52970 MW; 21E04F749844D88F CRC64;

Query Match 15.7%; Score 430; DB 6; Length 467;
Best Local Similarity 27.3%; Pred. No. 5.7e-22;
Matches 141; Conservative 72; Mismatches 203; Indels 94; Gaps 20;

QY 1 MALAPERAAPRVLFGEWLLGEISGCGYEGLOWLDEARTCFRVPWKHFAR-KDLSADARI 59
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 1 MAHPR-----RVRLKPWLVAQVDSGLYPLGLIWLHRDSKRFQIPWKHATRHSPQEEENTI 56
QY 60 FKAWAVAGRWPPSSRGGGPPPEAEATRAAGKTNFRCAIRSTRFVLMRDNSSG-PADP 118
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 57 FKAWAVETGKYQEG-----QEGVDDPDAK-----WKAQIRCALNKSREFNLMYDGTKEVPMNP 107
QY 119 HKVYALSRELWREGPGTDTQTEAEAPAAVPPQGGPPGPFLLAHAGLQAPGLP----- 173
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 108 VKIY-----QVC-----DIPQGGSIINP-----GSTGSAFWDEKD 138
QY 174 --APAGDGDLLQVQOCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGGLP 231
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 139 NDVDEDEDELDQSHVPIQDTFFFLNINGSPI-----APGS-----VGNCSVGNCS 188
QY 232 AGELYGWAVETTPSGQPPAALTTGEAAAPSPHQAEPYLSPSACTAVQPSGALDV 291
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 189 EAV-----W-----PKTEPLEMEVPQ--APIQPFYSSPELWISSLEMT-----DLDI 228
QY 292 TIMYKGRVLOKV-VGHP-SCTFLYGGPPD-----AVRATDPQVAPPSPAPLDPQKQ 343
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 299 KFQYRGKMGQTMVSNPQGCRLFGDLGPNQEELEFGPVSLEQVKGPPGPHITNEKQ 288
QY 344 RYTELLRHVAPGLHLELRGPMWARRMGKCKVYVEGPPGSGASPSPTPACLLPRNCDT 403
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 289 LFTSKLLDVMRGILLEVSHAIYAIRLCQCKVYWGSCAPSLVAPN-----LIERQKKVK 344
QY 404 IPDFRVFQELVEFARQRGRSPRYTIYLGFGQDLSAGRPKESILVVKLEPWLRCVHL 463
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 345 LFCLETFSLDLSHQKGIITQKPPPEIYLGFGEEWPDGKPKERKLIIVQIIPVAGMIYE 404
QY 464 GTQREGVSSLDLCLSSANSYDDIECFMEL 499
DQ ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DQ 405 MFGDFTSFSGSIRLQISTPD-IKDNIVAQKQL 439

RESULT 6
Q91VD0 PRELIMINARY; PRT; 467 AA.
AC Q91VD0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE BM282D4.4 (interferon regulatory factor 6).
GN IRF6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tromans A.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL365322; CAC42184.1; -.
DR EMBL; BC008515; AAH08515.1; -.
DR MGD; MGI:1859211; Irfg.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR PROSITE; PS00601; IRF; UNKNOWN 1.
SQ SEQUENCE 467 AA; 53110 MW; 1C564BC8D79C5259 CRC64;

Query Match 15.7%; Score 428; DB 11; Length 467;
Best Local Similarity 27.5%; Pred. No. 7.8e-22;
```

```
Matches 141; Conservative 72; Mismatches 212; Indels 88; Gaps 17;
QY 1 MALAPERARVLFGEWLLGEISGCGYGLQWLDDEARTCFRVPWKHFAR-KDLSEADARI 59
DB 1 MALHPR-----RVRLKMWLVQVDSGLYGLIWLHRDSKRFOIPWKHATRHSPOOEENTI 56
QY 60 FKAWAVARGRPSSGGGPPPEAETAEAGWKTNFRCLALRRFRFVMLRDN SGD-PADP 118
DB 57 FKAWAVETGKY---QEGVDDPDPAK-----WKAQLRCLANKSRFNLMDYDGTKEVPMNP 107
QY 119 HKYVALSRELWCWREGPGTQTEAEAPAAVPPPOGGPPGPFPLAHTHAGLQAPGPPAPAGD 178
DB 108 VKIY---QVC-----DIPQTQ-----SVINPGSTGSAFWD 135
QY 179 KGLLLQVQOCLADHLLTASGADPVPTKAPGEGEGLPLTGA-----TGACAGGPG 234
DB 136 EKONDVEDEE-----EDELQSHHVPIQDTFPLNINGSPIAPASVGNCSVGNCS 188
QY 235 LYGWAVETTPSPGPQPAALTTGEAAAPESPHQAEPLSPSPSACTAVQEPSGALDVTIM 294
DB 189 -----ESVMPKTEPLEMEVPOQAIOPFYSSPELWISLPMW-----DLDIKFQ 231
QY 295 YKGRVTLQKV-VGHP-SCTFLYGPDP-----AVRATDPOQVAFPSPAELPDOKQLRYT 346
DB 232 YRGKEYGQMTVNSNPOGCRFLFYGLGPMPODEELFGPVSLEQVKFGPEHITNEKQKLT 291
QY 347 EELLRHVAPGLHLELRGPQWARMGKCKYVWEGGPPGSPASSTPACLLPRNCDTPIED 406
DB 292 SKLVDVMDRGLLEVSGHAIYAIRLCKQVYWGSPCAPSLVAPN-----LIERQKVKLFC 347
QY 407 FRVFFOELVFRARQRGSPRYTIYLFQGDLSAGRPKESLVLVKLEPWLRCVHLEGTQ 466
DB 348 LETFSELIAHQKQIEKQEPFIYLCFGEWPDGKPLERKLIIVQVIVPVARMIYEMFS 407
QY 467 REGVSSLDSDLDLCLSSANSLYDDIECFMEL 499
DB 408 GDFTRSPDGSVRLQISTPD-IKDNIVAQLKQL 439

RESULT 7
Q8WNQ4 PRELIMINARY; PRT; 467 AA.
AC Q8WNQ4;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DE Interferon regulatory factor 6.
GN IRF6.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ILEUM;
RA Farber C.R., Raney N.E., Ernst C.W.;
RT "Characterization of the porcine IRF6 gene: cDNA cloning, expression
analysis and chromosomal localization."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327368; AAL37429.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR PRINTS; PR00267; INTERREGCT.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; UNKNOWN_1.
SQ SEQUENCE 467 AA; 53047 MW; 4AB757DA8013A3C2 CRC64;

Query Match 15.4%; Score 420; DB 6; Length 467;
Best Local Similarity 27.4%; Pred. No. 2.8e-21;
Matches 138; Conservative 71; Mismatches 210; Indels 84; Gaps 19;
QY 11 RVLFGEWLLGEISGCGYGLQWLDDEARTCFRVPWKHFAR-KDLSEADARIFKAWAVARGR 69

Matches 141; Conservative 72; Mismatches 212; Indels 88; Gaps 17;
QY 1 MALAPERARVLFGEWLLGEISGCGYGLQWLDDEARTCFRVPWKHFAR-KDLSEADARI 59
DB 1 MALHPR-----RVRLKMWLVQVDSGLYGLIWLHRDSKRFOIPWKHATRHSPOOEENTI 56
QY 60 FKAWAVARGRPSSGGGPPPEAETAEAGWKTNFRCLALRRFRFVMLRDN SGD-PADP 118
DB 57 FKAWAVETGKY---QEGVDDPDPAK-----WKAQLRCLANKSRFNLMDYDGTKEVPMNP 107
QY 119 HKYVALSRELWCWREGPGTQTEAEAPAAVPPPOGGPPGPFPLAHTHAGLQAPGPPAPAGD 178
DB 108 VKIY---QVC-----DIPQTQ-----SVINPGSTGSAFWD 135
QY 179 KGLLLQVQOCLADHLLTASGADPVPTKAPGEGEGLPLTGA-----TGACAGGPG 234
DB 136 EKONDVEDEE-----EDELQSHHVPIQDTFPLNINGSPIAPASVGNCSVGNCS 188
QY 235 LYGWAVETTPSPGPQPAALTTGEAAAPESPHQAEPLSPSPSACTAVQEPSGALDVTIM 294
DB 189 -----ESVMPKTEPLEMEVPOQAIOPFYSSPELWISLPMW-----DLDIKFQ 231
QY 295 YKGRVTLQKV-VGHP-SCTFLYGPDP-----AVRATDPOQVAFPSPAELPDOKQLRYT 346
DB 232 YRGKEYGQMTVNSNPOGCRFLFYGLGPMPODEELFGPVSLEQVKFGPEHITNEKQKLT 291
QY 347 EELLRHVAPGLHLELRGPQWARMGKCKYVWEGGPPGSPASSTPACLLPRNCDTPIED 406
DB 292 SKLVDVMDRGLLEVSGHAIYAIRLCKQVYWGSPCAPSLVAPN-----LIERQKVKLFC 347
QY 407 FRVFFOELVFRARQRGSPRYTIYLFQGDLSAGRPKESLVLVKLEPWLRCVHLEGTQ 466
DB 348 LETFSELIAHQKQIEKQEPFIYLCFGEWPDGKPLERKLIIVQVIVPVARMIYEMFS 407
QY 467 REGVSSLDSDLDLCLSSANSLYDDIECFMEL 499
DB 408 GDFTRSPDGSVRLQISTPD-IKDNIVAQLKQL 439

RESULT 8
Q96GL3 PRELIMINARY; PRT; 452 AA.
AC Q96GL3;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Similar to interferon regulatory factor 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009395; AAH09395.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR PROSITE; PS00601; IRF; UNKNOWN_1.
SQ SEQUENCE 452 AA; 49120 MW; 96B059A028AE6B64 CRC64;

Query Match 13.3%; Score 364.5; DB 4; Length 452;
Best Local Similarity 27.4%; Pred. No. 1.8e-17;
Matches 113; Conservative 43; Mismatches 153; Indels 103; Gaps 16;
QY 10 PRVLFGWLLGEISGCGYGLQWLDDEARTCFRVPWKHFARKDLSEADARIFKAWAVARGR 69
DB 6 PRIL--PWLVSQDLGLQLEGVAVWVNSRTRFRIPWKHGLRQDAQQEDFGIFQAWAEATGA 63
QY 70 WPSSRGSGGPPPEAETAEAGWKTNFRCLALRRFRFVMLRDN SGD-PADPDKVLSREL 129
DB 64 YVP-----GRDKDLPDT-----WKNFRSALNRKEGLRLAEDRSKDPHPKHIFEV---- 110
QY 130 WRSGPGTQTEAEAPAAVPPPOGGPPGPFPLAHTHAGLQAPGPPAPAGDKGDLQLQAVQ 189
```

```
Db 111 -NSGVG-----DFSQDTSPTDNTGG-----GSTSD-----TQ 136
QY 190 SCLADHLLTASGADPVPTKAPGEGQGLPLTACACAGGPGGLPAGELYGWAVETTPSPGPQ 249
Db 137 EDILDELL-GNNVLAPL-----DGGPP-----SLAVAPPCPQ 169
QY 250 PAALTGTGAAAPESPHQAEPPYLSPPSACTAVOEP-SPGALDVTIMYKGRTVLQKVGHP 308
Db 170 PLRSPSLDNPTP-----PPNLGSPENPLKRLLVPGEEWEFEVTAFYRGVFOQTISCP 223
QY 309 SCTFLYGGPDPAVRATDPOQVAFPSA-ELPDQKQLRYTEELLRHVAPGLHLELGRPOLW 367
Db 224 EGLRIVG-SEVGRITLPGWPTVLTPDGMSLTDRGWSYVRHVLSCLGGGLALWRAGQWLW 282
QY 368 ARRMGCKVYEVG-----GPPGASPTPACLAIPRNCDTPIFDPRVF 410
Db 283 AQRLGHCHTYWAVSELLPNSGHGPDG-----VPKDEGGVFDLGGP 325

RESULT 9
Q98TX7
ID Q98TX7 PRELIMINARY; PRT; 445 AA.
AC Q98TX7
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Interferon regulatory factor 4.
GN IRF-4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC STRAIN=SPAFAS; TISSUE=SPLEEN;
RX MEDLINE=21424622; PubMed=11533227;
RA Hrdlickova R., Nehyba J., Bose H.R. Jr.;
RT "Interferon Regulatory Factor 4 Contributes to Transformation of v-
Rel-Expressing Fibroblasts.";
RL Mol. Cell. Biol. 21:6369-6386(2001).
DR EMBL; AF320331; AAK08198.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERNEGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 445 AA; 51120 MW; 7CAE7BDF96780432 CRC64;

Query Match 13.3%; Score 364; DB 13; Length 445;
Best Local Similarity 26.4%; Pred. No. 1.9e-17;
Matches 130; Conservative 71; Mismatches 201; Indels 90; Gaps 20;

QY 16 EWLLEISSGCGEGQLWLDEARTCFRPWKHFARKDLS-EADARIFKAWAVARGWRPSS 74
Db 22 QWLIDQIDSGKYPLGWENDEKSI FRIPWKHAGQDYNEEDAALFKAWALFKGKF---- 77

QY 75 RGGGPPPEAETAEAGWKTNFRCAIARSTRFRVWLKDNNG-DPADPHKYVALSRELCLRSG 133
Db 78 REGIDKPDPT-----WKTRLCALNKSNDFFELVERSOLDISDPYKYVRIVPEGA-KKG 131

QY 134 PGTDQTEAEPAAVPPQGGPGPFLAHTHAGLQAPGLP-----APAGDKDLLLL 184
Db 132 AKONGMEEQPLMNHPPFITSP-----YTSLPQSQVNYVPHERNWREFAPEQHPDIPY 185

QY 185 QAVQSQSLADHLLTASGADPVPTKAPGEGQGLPLTACACAGGPGGLPAGELYGWAVETTP 244
Db 186 QCAS-----VFFAARGHHWQPGCENGCCQ-----VTGTFVACAPPESQ 223

QY 245 SPG-PQPAALTTGEEAAAPESPHQAEPPYLSPPSACTAVQEPSPGALDVTIMYKGRTVLQK 303
```

```
Db 224 TPGIPBESIRSGEALA-----LSDC-----RLHICLYIREMLVKEV 260
QY 304 VVGHF-SCTFLYGGPDPAVRATDPOQVAFPSAELPDQKQLRYTEELLRHVAPGLHLELR 362.
Db 261 TTSSPEGCRISQ-----QSVEVSSLEQVIFPYE---EDNSQRKNIEKLLSHLGRGVILWMA 314
QY 363 GPQLWARMGCKVYEVGPGPSAS--PSTPACLLPRNCDTPIFDPRVFFOELVEFRAR 420
Db 315 PDGLYAKRLCSRIYWD--GPLALCSDRPNK-----LERDQTKLFDTOQLAELQAF-AH 367
QY 421 QRRGSPRTIYLGFGQDLSAGRPKEKSLVLKLEPWLRCRVHLEGTQREGVSILDSDDL 480
Db 368 HGRPLRPTQVLCFGEFPDPQ-RORKLITAHVEPMFAQLYFFAQNSGCHLLRGVDLPE 426
QY 481 CLSSANSLYDDI 492
Db 427 LMTSPEDYHRSI 438

RESULT 10
Q98TX6
ID Q98TX6 PRELIMINARY; PRT; 409 AA.
AC Q98TX6
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Interferon regulatory factor 4 delta86.
GN IRF-4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC STRAIN=SPAFAS; TISSUE=SPLEEN;
RX MEDLINE=21424622; PubMed=11533227;
RA Hrdlickova R., Nehyba J., Bose H.R. Jr.;
RT "Interferon Regulatory Factor 4 Contributes to Transformation of v-
Rel-Expressing Fibroblasts.";
RL Mol. Cell. Biol. 21:6369-6386(2001).
DR EMBL; AF320332; AAK08199.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERNEGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 409 AA; 47489 MW; 0187B3782B62480E CRC64;

Query Match 13.3%; Score 362; DB 13; Length 409;
Best Local Similarity 25.0%; Pred. No. 2.4e-17;
Matches 123; Conservative 65; Mismatches 178; Indels 126; Gaps 18;

QY 16 EWLLEISSGCGEGQLWLDEARTCFRPWKHFARKDLS-EADARIFKAWAVARGWRPSS 74
Db 22 QWLIDQIDSGKYPLGWENDEKSI FRIPWKHAGQDYNEEDAALFKAWALFKGKF---- 77

QY 75 RGGGPPPEAETAEAGWKTNFRCAIARSTRFRVWLKDNNG-DPADPHKYVALSRELCLRSG 133
Db 78 REGIDKPDPT-----WKTRLCALNKSNDFFELVERSOLDISDPYKYVRI----- 123

QY 134 PGTDQTEAEPAAVPPQGGPGPFLAHTHAGLQAPGLPAPAGDKDLLLLQAVQSQCLA 193
Db 124 -----VPEGAKKGAKQNSMBEQPLM 143

QY 194 DHLTLASGADPVPTKAPGEGQGLPLTACACAGGPGGLPAGELYGWAVETTPSPGPQPAAL 253
Db 144 NH-----PPFITSP-----YTSLPQSQVNYM 164

QY 254 TTGE-----AAAPESPHQAEPPYLSPPS-PSACTAVQEPSPGA-----LDVTIMYKGRTVLQK 303
```

```
Db 165 VPHRNWREAPQPHDIPYQCASVFAAAGHHWGGPCENCRHLHICLYYREMLVKEV 224
Qy 304 VGHHP--SCTFLYGGPPDPAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLHLELR 362
Db 225 TTSPPGCRISQ---QSYEVSLQVIFYP---EDNSQKNIEKLLSHLGRGVILWMA 278
Qy 363 GPOLMARMKCKVYVEVGGPPGSAS--PSTPACLLPRNCDTPIFDPRVFFQBLVFRAR 420
Db 279 PDGLYAKRLCQSRIYWD--GPLALCSDRPNK---LERDQCKLFDQQLAELQAF-AH 331
Qy 421 QRGSPRYTYILGFGQDLNAGRKESLVILKLEPWLCHVLEGTQBEGVSSLDSDLDL 480
Db 332 HGRPLRYQVACLFGEEFPDQ--RORKLITAHVEPMPARQLYFYAQONGSHLLRGYDLPE 390
Qy 481 CLSSANSLYDDI 492
Db 391 LMTSPDYHRSI 402

RESULT 11
Q99419
ID Q99419 PRELIMINARY; PRT; 440 AA.
AC Q99419
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ICSAT transcription factor (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96239482; PubMed=8657101;
RA Yamagata T., Nishida J., Tanaka T., Sakai R., Mitani K., Yoshida M.,
RA Taniguchi T., Yazaki Y., Hirai H.;
RT "A novel interferon regulatory factor family transcription factor,
RT ICSAT/Pip/LSIRF, that negatively regulates the activity of interferon-
RT regulated genes.";
RL Mol. Cell. Biol. 16:1283-1294(1996).
DR EMBL; D78261; BAAL1335.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERFERGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
FT NON_TER
SQ SEQUENCE 440 AA; 49804 MW; 6A04159FAFBA2701 CRC64;

Query Match 12.6%; Score 344; DB 4; Length 440;
Best Local Similarity 28.1%; Pred. No. 4.6e-16;
Matches 124; Conservative 60; Mismatches 161; Indels 96; Gaps 21;

Qy 16 EWLIGISSCYEGLOWLDEARTCFRPWKHFARKDLS-EADARIFKAWAVARGRPPSS 74
Db 57 QWLIDQIDSKYPLGWENEKISFIRIPWKHAGKQDYNRDAAALFKAWLFGKF----- 112
Qy 75 RGGGPPPEATBAGKWNFRCALSTRFVWLNRNSG--DPADPHKVALSRELWCWREG 133
Db 113 REGIDKPDPT-----WKTRLCALNKSNDFEELVERSOLDISDPKYVRIIVEGAKK-- 165
Qy 134 PGTDQTEAEAP--AAVPPPGGPGPL-AHTHAGLQAPGLPAPAGDKGDLILLQAVQOS 190
Db 166 -GAKQLTLEDPQMSGHPTMTTPYSLPAQVNYMPP----- 203
Qy 191 CLADHLITASWGADPVPTKAPGEGQGLPLT-----GACAGGPGLPAGELYGWAV 240
Db 204 -----LDRSW-RDYVPDQPHPEIPYQCPMTFGRGHMGWPACENGCV-TGTFYACAP 255
Qy 241 EITTFSG-POPAALTGTGAAAAPSPHQAEPYLSFSPSACTAVQEPFGALDVTIMYKGR 299
```

```
Db 256 PESQAPGVTEPSIRSAAELA-----FSDC-----RLHICLYREIL 292
Qy 300 VLQKVVGHP--SCTFLYGGPPDPAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLH 358
Db 293 VKELTSSPEGCRISG---HTYDASNLQVLFYP---EDNCQKNIEKLLSHLGRGVV 346
Qy 359 LEIRGQPLWARMKCKVYVEVGGPPGSAS--PSTPACLLPRNCDTPIFDPRVFFQBLVE 416
Db 347 LWMAPDGLYAKRLCQSRIYWD--GPLALCNDRENK---LERDQCKLFDQQLAELQAF-AH 400
Qy 417 FRARQSGSPRYTYILGFGQD 437
Db 401 F-AHGRSLPRFQVTLFCGEE 420

RESULT 12
Q90W10
ID Q90W10 PRELIMINARY; PRT; 416 AA.
AC Q90W10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Interferon regulatory factor 10.
GN IRF10.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPAFAS; TISSUE=SPLEEN;
RA Nehyba J., Hrdlickova R., Bose H.R. Jr.;
RT "Interferon regulatory factor 10, a novel family member. IRF-10
RT expression is induced by interferons, concanavalin A, and the Rel/NF-
RT kappaB oncoprotein v-Rel.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF380350; AAK55444.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR ProDom; PD002355; IRF; 1.
DR PROSITE; PS00601; IRF; UNKNOWN 1.
SQ SEQUENCE 416 AA; 47646 MW; 669967280FEA967C CRC64;

Query Match 11.8%; Score 321; DB 13; Length 416;
Best Local Similarity 25.9%; Pred. No. 1.7e-14;
Matches 126; Conservative 65; Mismatches 161; Indels 134; Gaps 27;

Qy 16 EWLIGISSCYEGLOWLDEARTCFRPWKHFARKDL-SEADARIFKAWAVARGRPPSS 74
Db 12 EWLIAQIDSGRYPLGWENRERTLFRIPWKHAQKDYRQQAALFRAWAVYKGYHEGT 71
Qy 75 RGGGPPPEATBAGKWNFRCALSTRFVWLNRNSG--DPADPHKVALSRELWCWREG 133
Db 72 -----DKADPSTWTKRLCALNKSTDFQVPSRQIDISEPKYQI---VC--DG 117
Qy 134 PGTDQTEAEAPAAVPPPGGPGPLAHTHAGLQAPGLPAPAGDKGDLILLQAVQOSCLA 193
Db 118 TRDAEKDEKGRMQPTSSKDPQGHVAESHG-----TAG-----TC-- 154
Qy 194 DHLITASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGWAVETTPSPGQPAAL 253
Db 155 -HWPTL-----PLTAAHPAE--RGYHVRGI-----FYGW-----SP----- 182
Qy 254 TTGAAAAPSPHQAEPYLSFSPSACTAVQEPFGALDVTIMY-----KGRTV 300
Db 183 -TRSHLLPRAP--SFLPAEDVNHSDCW-----LHIRLYCDVLVKELTTRTAGCRI 231
Qy 301 LOKWVG--HPSCTFLYGGPPDPAVRATDPOQVAFPSPAELPDQKQLRYTEELLRHVAPGLH 358
Db 232 ASRTGPGYPSCM-----EQIEFPPPALGGGWTAVTVEVLERLLP--H 274
```

```

Db      297 KVVWSGPCAPSLAAPN-----LIERQKKVKLFCLETF 328

RESULT 14
Q924T6
ID      Q924T6      PRELIMINARY;      PRT;      215 AA.
AC      Q924T6;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      ISGF3G.
GN      ISGF3G.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=129/SVJ;
RX      MEDLINE=21243724; PubMed=11345588;
RA      Yawata M., Murata S., Tanaka K., Ishigatsubo Y., Kasahara M.;
RT      "Nucleotide sequence analysis of the ~35-kb segment containing
RT      interferon-gamma-inducible mouse proteasome activator genes.";
RL      Immunogenetics 53:119-129(2001).
RM      EMBL; AB053120; BAB47407.1; -.
RG      MGI; 107587; ISGF3g.
DR      InterPro; IPR001346; IRF.
DR      Pfam; PF00605; IRF; 1.
DR      ProDom; PD002355; IRF; 1.
DR      PROSITE; PS00601; IRF; UNKNOWN 1.
SQ      SEQUENCE 215 AA; 24028 MW; 5C2004D507F0C64D CRC64;

Query Match      7.1%; Score 194; DB 11; Length 215;
Best Local Similarity 35.5%; Pred. No. 4.7e-06;
Matches 39; Conservative 21; Mismatches 38; Indels 12; Gaps 3

QY      17 WLIGEISSGCGYGLQWLDEARTCFRPWKHPARKDLSE-ADARIFKAWAVAGRWPSSR 75
Db      15 WIVEQVESGHFPGVCWDAAKTWFIPIWKHAGKQDFREDQDAATPKKALPKK----- 68
QY      76 GGGPPPEAETAERAGWNTFRCAALRSTRFRVWLDRNSG-DRADPHKYVAL 124
Db      69 ----HKDGDGHGPAVWKTRLCALNKSFEFEVPERGMDVAEPYKVYRI 114

RESULT 15
Q9RKR9
ID      Q9RKR9      PRELIMINARY;      PRT;      1334 AA.
AC      Q9RKR9;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Putative multi-domain regulatory protein.
GN      SC02259 OR SCC75A.05C.
OS      Streptomyces coelicolor.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX      NCBI_TaxID=1902;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RA      Seeger K.J., Harris D.;
RL      Submitted (NOV-1999) to the EMBL/GenBank/DBSJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RA      Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL      Submitted (DEC-1999) to the EMBL/GenBank/DBSJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RA      Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,

```

RA Kinaishi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RN Mol. Microbiol. 21:77-96(1996).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL133220; CAB61705.1; -.
DR InterPro; IPR005158; BAD.
DR InterPro; IPR000767; Disease resist.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF03704; BAD; 1.
DR PRINTS; PR00364; DISEASERSIST.
DR PRINTS; PR01574; TUBBYPROTEIN.
DR ProDom; PD000329; Trans reg C; 1.
SQ SEQUENCE 1334 AA; 138787 MW; 78DC746883E8778C CRC64;

Query Match 6.7%; Score 183; DB 16; Length 1334;
Best Local Similarity 28.1%; Pred. No. 0.00021;
Matches 106; Conservative 25; Mismatches 150; Indels 96; Gaps 22;

Qy 18 LLGEISSGCGEGLQWLDEARTCFRVPWKHFARKDLSEADARIKAWAVAGRWPPSSRGG 77
Db 146 VLADLPDRTAEARW--ETRHFEALRARHTAALDLGOAHSLELTALCDGH----- 195

Qy 78 GPPPEAETAEER-----AGWKTNFRCALRSTRFRVMLRDNNG-DPA-----D 117
Db 196 -PLDEPLQALRLRALRDSGRTEALAAVEAVRR--LLADRLGTDPGPELTHAELLSPS 252

Qy 118 PHKYVALSRELKWREGP-----GTQTEAEAPAAVPPQGGP-PGFPLANTHAGL 166
Db 253 PTPTGCRSRTPGWTSGPGPASGAGAASGTDVASCAGAAAGDPDPASGPASGPAVAPGSGGG 312

Qy 167 QAPGLPAPAGDKDLLLQAVQOCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAG 226
Db 313 PAFGWWPAPGTAPGSGSTAPPHDTASAAD---TA---PAGFTSAPGTA---PAGTAAP 362

Qy 227 GPGL--PA-GELY--GWA--VETTPSPGQPOAALTGEEA-----AAPESPHOAEFYL 271
Db 363 APGTAGPAGTSYAPGTAPVAGTTPAGTAPAGTAPAGTAPAGTAPAGTAPVAGTTPAGT 422

Qy 272 SPSPSACTAV-----QEPSGALDVTIMYKGRTVLQKVGHPSCTFLYGPDPDAVRAT 324
Db 423 APAPGSPAPGCTVPAGTAPAGPQPA-----DGR---RPVTGPASGT---GPG-----AAT 468

Qy 325 DPQOVA-----FESPA 335
Db 469 PPEAAAAAASAGSPSPA 485

Search completed: June 18, 2003, 12:47:49
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 16:45:44 ; Search time 180 Seconds
(without alignments)
1430.977 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731

Sequence: 1 MALAPERAPRVLFGEWLLG.....SANSYDDIECFMLEQPA 503

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2715	99.4	503	1	IRF7_HUMAN	Q92985 homo sapien
2	1562	57.2	457	1	IRF7_MOUSE	P70434 mus musculus
3	1116	40.9	210	2	Q9UE79	Q9ue79 homo sapien
4	780	28.6	491	1	IRF3_CHICK	Q90643 gallus gall
5	778	28.5	491	2	Q90ZD4	Q90zd4 gallus gall
6	603	22.1	128	2	Q9HB64	Q9hb64 homo sapien
7	520	19.0	423	2	Q6PE40	Q6pe40 brachydanio
8	515.5	18.9	421	2	Q802H9	Q802h9 carassius a
9	487.5	17.9	498	1	IRF3_HUMAN	Q13568 homo sapien
10	484.5	17.7	488	2	Q6RCM8	Q6rcm8 homo sapien
11	482.5	17.7	514	2	Q64GB2	Q64gb2 homo sapien
12	458	16.8	427	1	IRF3_HUMAN	Q14653 homo sapien
13	453.5	16.6	419	2	Q764M6	Q764m6 sus scrofa
14	448	16.4	466	2	Q72XG6	Q72xg6 xenopus lae
15	445.5	16.3	492	2	Q6PCZ7	Q6pcz7 brachydanio
16	444.5	16.3	497	1	IRF5_MOUSE	P56477 mus musculus
17	440	16.1	459	2	O5Y578	O5y578 xenopus lae
18	439.5	16.1	382	2	Q6NZY9	Q6nzy9 brachydanio
19	438	16.0	419	1	IRF3_MOUSE	P70671 mus musculus
20	436.5	16.0	460	2	Q6GR23	Q6gr23 xenopus lae
21	436	16.0	467	1	IRF6_MOUSE	P97431 mus musculus
22	431	15.8	467	1	IRF6_HUMAN	O14896 homo sapien
23	430	15.7	467	2	Q9N136	Q9n136 ovis aries
24	428	15.7	467	2	Q91VD0	Q91vd0 mus musculus
25	424	15.5	517	2	Q63ZP8	Q63zp8 xenopus lae
26	422.5	15.5	460	2	Q6DD26	Q6dd26 xenopus lae
27	420	15.4	467	2	Q8WNQ4	Q8wnq4 sus scrofa
28	398	14.6	413	2	Q64GA9	Q64ga9 homo sapien
29	376	13.8	450	1	IRF4_MOUSE	Q64287 mus musculus
30	368	13.5	423	2	Q6DHK7	Q6dhk7 brachydanio
31	364.5	13.3	452	2	Q96GL3	Q96gl3 homo sapien

32	364.5	13.3	452	2	Q7Z5G6	Q7z5g6 homo sapien
33	364	13.3	445	2	Q98TX7	Q98tx7 gallus gall
34	363	13.3	451	1	IRF4_HUMAN	Q15306 homo sapien
35	362	13.3	409	2	Q98TX6	Q98tx6 gallus gall
36	344	12.6	440	2	Q99419	Q99419 homo sapien
37	338	12.4	425	1	ICSB_CHICK	Q90871 gallus gall
38	333.5	12.2	424	1	ICSB_MOUSE	P23611 mus musculus
39	331	12.1	392	2	Q6NY63	Q6ny63 brachydanio
40	328.5	12.0	393	1	IRTF_HUMAN	Q00378 homo sapien
41	322	11.8	397	2	Q6NSI6	Q6nsi6 xenopus lae
42	321	11.8	416	2	Q90W10	Q90w10 gallus gall
43	318.5	11.7	389	2	Q64GB0	Q64gb0 homo sapien
44	308.5	11.3	426	1	ICSB_HUMAN	Q02556 homo sapien
45	306	11.2	398	2	Q68FP4	Q68fp4 rattus norv

ALIGNMENTS

RESULT 1
ID IRF7_HUMAN STANDARD; PRT; 503 AA.
AC Q92985; O00331; O00332; O00333; O75924;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Interferon regulatory factor 7 (IRF-7).
GN Name=IRF7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Spleen;
RA Grossman A., Nicholl J., Antonio L., Luethy R., Suggs S.,
RA Sutherland G.R., Mak T.W.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RX MEDLINE=97459673; PubMed=9315633;
RA Zhang L., Pagano J.S.;
RT "IRF-7, a new interferon regulatory factor associated with Epstein-Barr virus latency.";
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM D).
RX MEDLINE=99003279; PubMed=9786932; DOI=10.1074/jbc.273.44.29210;
RA Au W.-C., Moore P.A., Lafleur D.W., Tombal B., Pitha P.M.;
RT "Characterization of the interferon regulatory factor-7 and its potential role in the transcription activation of interferon A genes.";
RN [4]
RP J. Biol. Chem. 273:29210-29217(1998).
RN [4]
RP SPLICED ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
RA Hillman R.T., Green R.E., Brenner S.E.;
RT "An unappreciated role for RNA surveillance.";
RN [5]
RP Genome Biol. 5:RSEARCH008.1-RESEARCH008.16(2004).
CC -!- FUNCTION: Transcriptional activator. Binds to the interferon-promoter (Op) of EBV nuclear antigen-1 (EBNA1).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=A;
CC IsoId=Q92985-1; Sequence=Displayed;
CC Name=B; Synonyms=Beta;
CC IsoId=Q92985-2; Sequence=VSP_002760;
CC Name=C; Synonyms=Gamma;
CC IsoId=Q92985-3; Sequence=VSP_002758, VSP_002759;
CC Notes=May be produced at very low levels due to a premature stop codon in the mRNA, leading to nonsense-mediated mRNA decay;

CC CC Names=D; Synonyms=N;
 CC CC ISOId=Q92985-4; Sequence=VSP_002757;
 CC CC -I- TISSUE SPECIFICITY: Expressed predominantly in spleen, thymus and
 CC CC peripheral blood leukocytes
 CC CC -I- SIMILARITY: Belongs to the IRF family.
 CC CC
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC
 CC CC ENBL; U73036; AAB17190.1; -;
 CC CC ENBL; U53830; AAB80686.1; -;
 CC CC ENBL; U53831; AAB80688.1; -;
 CC CC ENBL; U53832; AAB80690.1; -;
 CC CC ENBL; AF076494; AAC70999.1; -;
 CC CC HSSP; P23906; 2IRF.
 CC CC TRANSFAC; T04674; -;
 CC CC TRANSFAC; T05106; -;
 CC CC Genew; HGNC:6122; IRF7.
 CC CC MIM; 605047; -;
 CC CC GO; GO:0005737; C:cytoplasm; TAS.
 CC CC GO; GO:0005634; C:nucleus; ISS.
 CC CC GO; GO:0003677; F:DNA binding; ISS.
 CC CC GO; GO:0003704; F:specific RNA polymerase II transcription fa...; TAS.
 CC CC GO; GO:0006954; P:inflammatory response; ISS.
 CC CC GO; GO:0001022; P:negative regulation of transcription from P...; TAS.
 CC CC GO; GO:0046731; P:passive viral induction of host immune resp...; ISS.
 CC CC GO; GO:0006974; P:response to DNA damage stimulus; ISS.
 CC CC GO; GO:000615; P:response to virus; ISS.
 CC CC GO; GO:0006367; P:transcription initiation from Pol II promoter; ISS.
 CC CC InterPro; IPR001346; IRF.
 CC CC InterPro; IPR008984; SMAD FHA.
 CC CC InterPro; IPR009058; Wing_hlx_DNA_bnd.
 CC CC Pfam; PF00605; IRF.1.
 CC CC PRINTS; PR00267; INTERFERGCT.
 CC CC PRODOM; PD002355; IRF.1.
 CC CC SMART; SM00348; IRF.1.
 CC CC PROSITE; PS00601; IRF.1.
 CC CC Activator; Alternative splicing; DNA-binding; Nuclear protein;
 CC CC Transcription regulation.
 CC CC DNA_BIND 13 122 Tryptophan pentad repeat.
 CC CC VARSPPLIC 1 6 MALAPE -> MPVPERPAAGDPSRPRCTR (in isoform
 CC CC D).
 CC CC /FTId=VSP_002757.
 CC CC GGGPGGFLAHTHA -> AQSLLGSGCTGGQ (in
 CC CC isoform C).
 CC CC /FTId=VSP_002758.
 CC CC Missing (in isoform C).
 CC CC /FTId=VSP_002759.
 CC CC Missing (in isoform B).
 CC CC /FTId=VSP_002760.
 CC CC E -> K (in Ref. 2).
 CC CC Q -> R (in Ref. 3).
 CC CC SEQUENCE 503 AA; 54278 MW; AA6A39E0E272727C CRC64;
 CC CC
 CC CC Query Match 99.4%; Score 2715; DB 1; Length 503;
 CC CC Best Local Similarity 99.4%; Pred. No. 1.8e-142;
 CC CC Matches 500; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC CC
 CC CC 1 MALAPEAARVLFGEWLLGEISGCGVEGLQWLEARTCFRPWKHFARKDLSEADARIF 60
 CC CC |||||
 CC CC 1 MALAPEAARVLFGEWLLGEISGCGVEGLQWLEARTCFRPWKHFARKDLSEADARIF 60
 CC CC |||||
 CC CC 61 KANAVARGRPSPSRGGPPPEAETAEAGWKINFRCALSTRFRVFLRNDSGDPADPHK 120
 CC CC |||||
 CC CC 61 KANAVARGRPSPSRGGPPPEAETAEAGWKINFRCALSTRFRVFLRNDSGDPADPHK 120
 CC CC |||||
 CC CC 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPPGGPGPFLLAHTHAGLQAPGFLPAPAGDEG 180
 CC CC |||||

Db 121 VYALSRELCWREGPGTDTQTEAEAPAAVPPPGGPGPFLLAHTHAGLQAPGFLPAPAGDEG 180
 QY 181 DLLLQAVQOSCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGNVAV 240
 Db 181 DLLLQAVQOSCLADHLLTASWGADPVPTKAPGEGQGLPLTGACAGGPGLPAGELYGNVAV 240
 QY 241 ETTTPSGPQPAALTTCGAAAPESPHQAEPLYSPSPSACTAVQSPSCALDVTIMYKRTV 300
 Db 241 ETTTPSGPQPAALTTCGAAAPESPHQAEPLYSPSPSACTAVQSPSCALDVTIMYKRTV 300
 QY 301 LQKVGHPCSTFLYGGPDPAVRATDPOQVAFSPSAPLPOKQRYTEELLRHVAPGLHLE 360
 Db 301 LQKVGHPCSTFLYGGPDPAVRATDPOQVAFSPSAPLPOKQRYTEELLRHVAPGLHLE 360
 QY 361 LRGPQLWARRMGKCKVYWEVGGPPGSGASPSPTACLLPRNCDTPIFDFRVFFQELVEFRAR 420
 Db 361 LRGPQLWARRMGKCKVYWEVGGPPGSGASPSPTACLLPRNCDTPIFDFRVFFQELVEFRAR 420
 QY 421 QRRGSPRYTYILGFGQDLSAGRPKEKSLVLKLEPWLRCRVHLEGTQREGVSSLDSDLDL 480
 Db 421 QRRGSPRYTYILGFGQDLSAGRPKEKSLVLKLEPWLRCRVHLEGTQREGVSSLDSDLSL 480
 QY 481 CLSSANSLYDDIECFLEMELEOPA 503
 Db 481 CLSSANSLYDDIECFLEMELEOPA 503
 CC CC
 CC CC RESULT 2
 CC CC IRF7_MOUSE STANDARD; PRT; 457 AA.
 CC CC AC P70434;
 CC CC DT 01-NOV-1997 (Rel. 35, Created)
 CC CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC CC DT 05-JUL-2004 (Rel. 44, Last annotation update)
 CC CC DE Interferon regulatory factor 7 (IRF-7).
 CC CC GN Name:IRF7;
 CC CC OS Mus musculus (Mouse).
 CC CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC CC OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC CC OX NCBI_TaxID=10090;
 CC CC RN [1]
 CC CC RP SEQUENCE FROM N.A.
 CC CC RC STRAIN=BALB/c; TISSUE=Spleen;
 CC CC RA Grossman A., Nicholl J., Antonio L., Luethy R., Suggs S.,
 CC CC RA Sutherland G.R., Mak T.W.;
 CC CC RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC CC CC -I- FUNCTION: transcriptional activator. Binds to the interferon-
 CC CC stimulated response element (ISRE) in IFN promoters and in the Q
 CC CC promoter (Qp) of EBV nuclear antigen-1 (EBNA1) (By similarity).
 CC CC -I- SUBCELLULAR LOCATION: Nuclear (potential).
 CC CC -I- SIMILARITY: Belongs to the IRF family.
 CC CC
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC
 CC CC ENBL; U73037; AAB18626.1; -;
 CC CC DR HSSP; P23906; 1IRF.
 CC CC DR MGD; MGI:1859212; IRF7.
 CC CC DR InterPro; IPR001346; IRF.
 CC CC DR InterPro; IPR008984; SMAD FHA.
 CC CC DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
 CC CC DR Pfam; PF00605; IRF.1.
 CC CC DR PRINTS; PR00267; INTERFERGCT.
 CC CC DR PRODOM; PD002355; IRF.1.
 CC CC DR SMART; SM00348; IRF.1.
 CC CC DR PROSITE; PS00601; IRF.1.
 CC CC KW Activator; DNA-binding; Nuclear protein; Transcription regulation.
 CC CC FT DNA_BIND 11 122 Tryptophan pentad repeat.
 CC CC FT

Qy 2 ALAPERAAPRVLFGEWLLGEISSGCGYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFK 61

```
Db 3 ALDSEGAQKLRGFWLLNVAVSSGLYRGLCWIDPRRIFRIPWKHNARKDVTSSDVEIFK 62
QY 62 AWAVARGRWPSSRGSGPPPEAETAEARAGKTNFRCALRSTRFRVLMRDNSSGDPADPHKV 121
Db 63 AWAKASGRY-----EGNAEDPAKWKTNFRCALRSTHFMFLLEDKSKNDPDHVK 111
QY 122 YALSRELWCWREGPGTQDTEAEAPAAVPPQ--GGP-----PGPFLAHTHAGLQ---- 167
Db 112 YAVA-----SGVENDRGSGGPVAGALQQOQPOLLLNHHDLALENTPT 152
QY 168 -AFGPLPAPAGDKGDL-LLQAQVOSCLADHLLTASWGADVPVTKAPGEG--QEGPLPTG- 222
Db 153 DSTEGLVAAALTCVDLQSLVQHCHNISAL-----GSQPTLWAHTGDALPEDALLPGQ 207
QY 223 -ACAGGFLPAGLYGW-AVETTPSPGPQ-----AALTTCGAAAPSPHQAE 269
Db 208 DGLCLPG-----QFQWRQLEBLLGNQPLTGGCGQDQAGALPVSECAIPAPSPAE 262
QY 270 YL-----SPSPSACTAVQEPSGALDVTIMYKGRVTLQKVGHPSCTFLYGPDPVAVRATD 325
Db 263 LLFQSANPAPPAGDIGGLPLLDITIIYRGMVYQEQVDDSRCLVLAQPLDPAV--AE 320
QY 326 PQQVAFPSPAELPDQQLRYTELLRHVAPGLHLELRGPOLWARRMKCKVYWEV----- 380
Db 321 QRLVLPSPASLPDPQRRTYEDLLE--VAGLRLEORAGQLLATRLKCKVFWALSQOLE 378
QY 381 GGPFGSASPSTPACLLPRNCDTPIFDPRVFFQELVEFRARQRRGSPRYTIIYLGFGDLSA 440
Db 379 GGEP-----PLNLLHRDQETIIFDRVFCTELDRFRDRRRERSPDFIIFLCGQCFSS 431
QY 441 GRPKESLVILVLEPWLRCVHLEGTQREGVSLDSSDLCLSSANSYLDIECFLEMELE 500
Db 432 TKPKESKILVLPVQCEVYVQVQRGASSLNGVSLQSLSDSFNLFELIEQYHMQTD 491

RESULT 5
Q90ZD4 PRELIMINARY; PRT; 491 AA.
AC Q90ZD4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Interferon regulatory factor-3.
GN Names=IRF-3;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA May D.L., Grant C.E., Deeley R.G.;
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268079; AAK58583.1; -.
DR HSSP; P15314; 1IF1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006055; IRF; 1.
DR PRINTS; PR00267; INTFRNREGFCT.
DR PRODOM; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 491 AA; 54440 MW; 90EC080A2624B1B4 CRC64;

Query Match 28.5%; Score 778; DB 2; Length 491;
Best Local Similarity 38.5%; Pred. No. 2.2e-35;
Matches 208; Conservative 59; Mismatches 181; Indels 92; Gaps 17;

QY 2 ALAPERAAPRVLEGMWLGISGCTYEGLOLWDEARTCFRVPWKHFARKDLSADARIFK 61
Db 3 ALDSEGAQKLRGFWLLNVAVSSGLYRGLCWIDPRRIFRIPWKHNARKDVTSSDVEIFK 62
```

```
QY 62 AWAVARGRWPSSRGSGPPPEAETAEARAGKTNFRCALRSTRFRVLMRDNSSGDPADPHKV 121
Db 63 AWAKASGRY-----EGNAEDPAKWKTNFRCALRSTHFMFLLEDKSKNDPDHVK 111
QY 122 YALSRELWCWREGPGTQDTEAEAPAAVPPQ--GGP-----PGPFLAHTHAGLQ---- 167
Db 112 YAVA-----SGVENDRGSGGPVAGALQQOQPOLLLNHHDLALENTPT 152
QY 168 -AFGPLPAPAGDKGDL-LLQAQVOSCLADHLLTASWGADVPVTKAPGEG--QEGPLPTG- 222
Db 153 DSTEGLVAAALTCVDLQSLVQHCHNISAL-----GSQPTLWAHTGDALPEDALLPGQ 207
QY 223 -ACAGGFLPAGLYGW-AVETTPSPGPQ-----AALTTCGAAAPSPHQAE 269
Db 208 DGLCLPG-----QFQWRQLEBLLGNQPLTGGCGQDQAGALPVSECAIPAPSPAE 262
QY 270 YL-----SPSPSACTAVQEPSGALDVTIMYKGRVTLQKVGHPSCTFLYGPDPVAVRATD 325
Db 263 LLFQSANPAPPAGDIGGLPLLDITIIYRGMVYQEQVDDSRCLVLAQPLDPAV--AE 320
QY 326 PQQVAFPSPAELPDQQLRYTELLRHVAPGLHLELRGPOLWARRMKCKVYWEV----- 380
Db 321 QRLVLPSPASLPDPQRRTYEDLLE--VAGLRLEORAGQLLATRLKCKVFWALSQOLE 378
QY 381 GGPFGSASPSTPACLLPRNCDTPIFDPRVFFQELVEFRARQRRGSPRYTIIYLGFGDLSA 440
Db 379 GGEP-----PLNLLHRDQETIIFDRVFCTELDRFRDRRRERSPDFIIFLCGQCFSS 431
QY 441 GRPKESLVILVLEPWLRCVHLEGTQREGVSLDSSDLCLSSANSYLDIECFLEMELE 500
Db 432 TKPKESKILVLPVQCEVYVQVQRGASSLNGVSLQSLSDSFNLFELIEQYHMQTD 491

RESULT 6
Q9HB64 PRELIMINARY; PRT; 128 AA.
AC Q9HB64;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Interferon regulatory factor-7H (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20493557; PubMed=10924517; DOI=10.1074/jbc.M005288200;
RA Lu R., Au W.-C., Yeow W.-S., Hageman N., Pitha P.M.;
RT "Regulation of the promoter activity of interferon regulatory factor-7
gene: ACTIVATION BY INTERFERON AND SILENCING BY HYPERMETHYLATION.";
RL J. Biol. Chem. 275:31805-31812(2000).
DR EMBL; AF271159; AAG30003.1; -.
DR HSSP; P23906; 2IRF.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003677; F:RNA binding; TAS.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; TAS.
DR GO; GO:0006954; P:inflammatory response; TAS.
DR GO; GO:0046731; P:passive viral induction of host immune resp. .; TAS.
DR GO; GO:0006974; P:response to DNA damage stimulus; NAS.
DR GO; GO:0006615; P:response to virus; TAS.
DR GO; GO:0006367; P:transcription initiation from Pol II promoter; TAS.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTFRNREGFCT.
DR PRODOM; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
FT NON_TER 128 AA; 14363 MW; 09B5B6F3CC11BB15 CRC64;
SQ SEQUENCE 128 AA; 14363 MW; 09B5B6F3CC11BB15 CRC64;
```

Query Match 22.1%; Score 603; DB 2; Length 128;
 Best Local Similarity 100.0%; Pred. No. 2.8e-26;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RAAPRVFGWLLGRISSCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVA 66
 DB 20 RAAPRVFGWLLGRISSCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVA 79

QY 67 RGRWPPSSRGSGGPPPEAETAEAGKTNFRCLALSTRFRVLMRLDNGSDP 115
 DB 80 RGRWPPSSRGSGGPPPEAETAEAGKTNFRCLALSTRFRVLMRLDNGSDP 128

RESULT 7
 Q6PE40
 ID Q6PB40 PRELIMINARY; PRT; 423 AA.
 AC Q6PB40;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DE Interferon regulatory factor 7.
 GN Name=irf7;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner B., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC058298; AAH58298.1; -
 DR ZFIN; ZDB-GENE-040426-1518; irf7.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001346; IRF.
 DR InterPro; IPR008984; SMAD FHA.
 DR Pfam; PF00605; IRF.1
 DR PRINTS; PR00267; INTERREGCT.
 DR ProDom; PD002355; IRF.1.
 DR SMART; SM00348; IRF.1.
 SQ SEQUENCE 423 AA; 49060 MW; 3A59DE6FB6D2275E CRC64;

Query Match 19.0%; Score 520; DB 2; Length 423;
 Best Local Similarity 30.9%; Pred. No. 3.5e-21;
 Matches 151; Conservative 67; Mismatches 174; Indels 96; Gaps 16;

QY 14 FGEWLLGRISSCYEGLOWLDEARTCFRVPWKHFARKDLSEADARIFKAWAVARGR---W 70
 DB 10 FGPWLLIEQVESQYEGLSMI--GHDFRIPWKHNARRDLGDADVKIFKEWAIYVSGKINEY 67

QY 71 PPSRRGGGPPPEAETAEAGKTNFRCLALSTRFRVLMRLDNGSDPADPHKVVALSRELWCW 130
 DB 68 P-----NDKAKWKTNFRCLALSHUKNFEMLEHDSKDPDDQHKIYRIIR----- 109

QY 131 REGPGTDTAEAPAAVPPPGGPPG-PFLAHTHAGLQAPGLPAPAG-----DKGDL 182
 DB 110 -----PQNHQEIQSAIOSAEAVQQLPFAEVY---NASNHSQDMLELNLVETMDL 160

QY 183 LIQAVQOCLADHLLLTASWGADPVFTKAPGEGQGLPLTGACAGGGLPAGBELYGHAVET 242
 DB 161 NLHAVSQS-----LXTYS-----QPNIQTTSSN-----YFET 187

QY 243 TSPGPGQPAALTTGEAAAPESPQAEPLVSPSPSACTAQQEPPGALDVTIMVKGRTVLQ 302
 DB 188 TVSDGP---CMQNNIPASVQOSH-----TTVDQWNLCDLEISINRTRTEVLK 231

QY 303 KVVGHPSCTFLYGGPPDPVAVRATDPOQVAFPPSPAEPLDQKQLRYTBELLRHVAPGLHLELR 362
 DB 232 TFLCSSLIHFFVQCDPSRLG---EQIRFPTTECLIDVKIQYTKRILDSIQRLGLEVN 288

QY 363 GPQLWARWGKCKVYVWVGPGPSASPTPACLLPRNCPTPIDFRVFFQELVEPRAROR 422
 DB 289 QYGIYGFQRDKCKVFVSTSDPCIEIQKPEPRK--LQNVKEQLLSFDKYIRDLDFK-ENR 345

QY 423 RGSPTYTYLGFQDLSAGRPKEKSLVLVLEPWLKRVHLEGTORGESVLSLSSDLDLCL 482
 DB 346 GGSPTYTYLGFQDLSAGRPKEKSLVLVLEPWLKRVHLEGTORGESVLSLSSDLDLCL 405

QY 483 SSANSLYD 490
 DB 406 SH-NSLYD 412

RESULT 8
 Q802H9
 ID Q802H9 PRELIMINARY; PRT; 421 AA.
 AC Q802H9;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Interferon regulatory factor 7.
 GN Name=IRF7;
 OS Carassius auratus (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Zhang Q., Xu D., Hu C., Gui J.;
 RT "Identification of antiviral-relevant genes in the cultured fish cells
 induced by inactivated virus.";
 RL Chin. Sci. Bull. 48:581-588(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22912602; PubMed=14550671; DOI=10.1016/S1050-4648(03)00025-1;
 RA Zhang Y., Hu C., Zhang J., Huang G., Wei L., Zhang Q., Gui J.;
 RT "Molecular cloning and characterization of crucian carp (Carassius
 auratus L.) interferon regulatory factor 7(1).";
 RL Fish Shellfish Immunol. 15:453-466(2003).
 DR EMBL; AY177629; RAO18646.1; -
 DR HSP; F15314; IIF1.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001346; IRF.
 DR InterPro; IPR008984; SMAD FHA.
 DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
 DR PRINTS; PR00267; INTERREGCT.
 DR ProDom; PD002355; IRF.1.
 DR SMART; SM00348; IRF.1.
 SQ SEQUENCE 423 AA; 49060 MW; 3A59DE6FB6D2275E CRC64;

```
DR Pfam: PF00605; IRF; 1.
DR PRINTS: PR00267; INTERREGCT.
DR PRODOM: PD002355; IRF; 1.
DR SMART: SM00348; IRF; 1.
SQ SEQUENCE 421 AA; 49056 MW; 605B43465FD89ABE CRC64;

Query Match 18.9%; Score 515.5; DB 2; Length 421;
Best Local Similarity 30.3%; Pred. No. 61e-21;
Matches 150; Conservative 70; Mismatches 160; Indels 115; Gaps 20;

QY 14 FGMWLLGEISSGCGYGLQWLDIARTCFRVPWKHFARKDLEADARIFKAWAVARGRWPPS 73
DQ |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 13 FGPWLLGEISSGCGYGLQWLDIARTCFRVPWKHFARKDLEADARIFKAWAVARGRWPPS 66
QY 74 SRGGGPPPEAETARAGWTKNFCALRSTRFRVWLRDNGSDPADPKVYALSRELWCREG 133
DQ |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 67 -----INHPNDKAKWTKNFCALYSLKFNEMLEHSDKDPDQKHVKVRIIR----- 112

QY 134 PGTDQTEAEAPAAVPPQGGPPGPFLLAHTHAGLQAPGLPAPAGDKGDLLOAVQOCSLA 193
DQ |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 113 ---PQNHQEIQSAEPVQL---PLPFTISEVY-----NNYMHEDMEQELLIS 150

QY 194 D-----HLLTASGADVPVTKAPGEGGLPLTGACAGGGLPAGELYGWAVETTPSPG-- 247
DQ |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 151 QVETMHLNQOS--AEPQWDC---SQNIQTTSRYFG-----TPYPEQC 190

QY 248 -----POPALTTCGAAPSPHQAEPYLSPS-----PSACTAVQEPSPGALDVTIMYKGR 298
DQ |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 191 MNMMPDPV-----QQPYTAQOWNVPALCD-----LEISINRYKT 226

QY 299 TVLQKVGVHPSCTFL---YGPDPVAVRATDPQVAFPSPAELPDQKQRYTEBLLRHVAP 355
DQ |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 227 EVLKTRL---CSSLVQFYQCPSSELRGEP--IRFPATEGLDHKQIQFTKILDSIOR 280

QY 356 GLHLELRGPOLWARMKCKKVVVEVGGPPGASPSPTACLLPNCOTPIPDFRVFFOELV 415
DQ |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 281 GLQLEVNQYGIYGRQCKKVFVSTSDPSEIQNPPEPK--LHQNRSRQLFSFKYIRDLM 338

QY 416 EFRARQRGSRPTIYILGFGQDLISAGRPKESLVLVKLEPWLRCVHLEGTQREGVSLDS 475
DQ |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 339 GFK-ENRGRSPDYIYLCFGKULPDGKPLEKLLTVKVPVLCRELHERAQMEGASLS-R 396

QY 476 SLDLCLSSANSYLD 490
DQ |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 397 GNVSLQISH-NSLFD 410

RESULT 9
IRF5 HUMAN STANDARD; PRT; 498 AA.
AC Q1358; O9B0F0.
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Interferon regulatory factor 5 (IRF-5).
GN Name=IRF5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grosseman A., Mittrucker H.W., Lantonio L., Mak T.W.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore A., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
```

```
QY 259 AAPESHQAEVYLSPPSPSACTAVQSPSGALDVTIMYKGRVTLQKVGHGHP--SCTFLYGGP 317
DB 228 --PPAGEQLPDLILSHPLMLPTD-----LEIKFYGRGPRPRALTISNPHGCRFLFYS-- 277
QY 318 DPAVATDQ-----OYAFPSPAELPDOKQLRYTEELLRHVAPGLHLELGPQLW 367
DB 278 --OLEATQOVELFGPISLEQVRFPSPEDIPSDKQRFYTNQLLDVLDRLGLILQLOQGDLY 335
QY 368 ARMGCKVYWEVGGPGSGASPTACLLPRNCDTPIFDRVFFQELVEFRARQRGSPR 427
DB 336 AIRLCOCKVFW--SGFCASAHSDCPN-PIQREVKTGLFSLHFLNELILFQKQTNTPPP 392
QY 428 YTYLFGQDLGAGRPKESLVLVKLEPWLCRVHLEGTQREGVSSLDSDLDL 480
DB 393 FEIFFCGEWPDRKPREKLIITVQVVPVAARLLEMFSGELSWADSIRLQI 445

RESULT 10
Q6RCM8 PRELIMINARY; PRT; 488 AA.
AC Q6RCM8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Interferon regulatory factor 5 variant 4 (Interferon regulatory factor
DE 5 variant 3).
GN Name=IRF5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21316468; PubMed=11303025; DOI=10.1074/jbc.M101216200;
RA Barnes B.J., Moore P.A., Pitha P.M.;
RT "Virus-specific activation of a novel interferon regulatory factor,
RT IRF-5, results in the induction of distinct interferon alpha genes.";
RL J. Biol. Chem. 276:23382-23390(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Barnes B.J., Pitha P.M.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY504947; AAR90326.1; -.
DR EMBL; AY504946; AAR90325.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR009058; wing_hlx_DNA_bnd.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERNRGFCF.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
SQ SEQUENCE 488 AA; 54942 MW; 4E43B54EE74A7E37 CRC64;
```

```
Query Match 17.7%; Score 484.5; DB 2; Length 488;
Best Local Similarity 28.6%; Pred. No. 3.6e-19;
Matches 151; Conservative 59; Mismatches 171; Indels 147; Gaps 20;
```

```
QY 1 MALAPERAAPRVLFGEWLGEISSGCEGLQWLDEARTCFRVPWKHFARKDLS-EADARI 59
DB 7 VAPTTPR---RVRLKPWLVAQVNSCYQPGLOWNGEKKLFCIPWRHATHRGPSQGDNTI 63
QY 60 FXAWAVARGWPPSSRGSGGPPPEATAERAGWKTNFRCALSTRFRVLMRLDSGD--PADP 118
DB 64 FRAWAKETGKYTEG-----VDEADPAKWANKRLCALNKSRRDPRLIYDGRDMPPOP 114
QY 119 HKVYALSRELCHWREGTGTQTEAEAPAAVPPQGGPPG-----PFLAHTH 163
DB 115 YKIY-----EVC-SNGF-----APTDSQPPEDYSFGAGEEEEEELQRLMPLSLSLTD 170
```

```
QY 149 PPOGPPP---GFLANHTAGLOAPGFLPAPAGDKGDLLOAVQOQCLADHLLTASWGADP 205
DB 171 PPTLQPPVVVLG-----PAP-----DP 187
QY 206 VPTKAPGEGQGLPLTGACAGGPGLPAG--ELYGWAVETTPSGPOPAAATTGEEAAAPES 263
DB 188 SPLAPP-----PGNPAGFRELLSEVLE-----PGPLPASL-----PPA 220
QY 264 PHQAEVYLSPPSACTAVQSPSGALDVTIMYKGRVTLQKVGHGHP--SCTFLYGGPDPVR 322
DB 221 GFQLLPDLILSHPLMLPTD-----LEIKFYGRGPRPRALTISNPHGCRFLFYS-----QLE 270
QY 323 ATDPO-----QVAFPSPAELPDOKQLRYTEELLRHVAPGLHLELGPQLWARRMG 372
DB 271 ATQEOVELFGPISLEQVRFPSPEDIPSDKQRFYTNQLLDVLDRLGLILQLOQGDLYAIRLC 330
QY 373 KCKVYWEVGGPGSGASPTACLLPRNCDTPIFDRVFFQELVEFRARQRGSPRTIYL 432
DB 331 QCKVFW--SGFCASAHSDCPN-PIQREVKTGLFSLHFLNELILFQKQTNTPPPPEIFF 387
QY 433 GFGQDLSAGRPKESLVLVKLEPWLCRVHLEGTQREGVSSLDSDLDL 480
DB 388 CFGEWPDREKPREKLIITVQVVPVAARLLEMFSGELSWADSIRLQI 435

RESULT 11
Q64GB2 PRELIMINARY; PRT; 514 AA.
AC Q64GB2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Interferon regulatory factor 5 variant 5.
GN Name=IRF5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hu G., Mancel M.E., Barnes B.J.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY693665; AAU12877.1; -.
SQ SEQUENCE 514 AA; 57770 MW; D554D121D9E35893 CRC64;
```

```
Query Match 17.7%; Score 482.5; DB 2; Length 514;
Best Local Similarity 29.6%; Pred. No. 4.9e-19;
Matches 154; Conservative 60; Mismatches 201; Indels 105; Gaps 19;
```

```
QY 1 MALAPERAAPRVLFGEWLGEISSGCEGLQWLDEARTCFRVPWKHFARKDLS-EADARI 59
DB 7 VAPTTPR---RVRLKPWLVAQVNSCYQPGLOWNGEKKLFCIPWRHATHRGPSQGDNTI 63
QY 60 FXAWAVARGWPPSSRGSGGPPPEATAERAGWKTNFRCALSTRFRVLMRLDSGD--PADP 118
DB 64 FRAWAKETGKYTEG-----VDEADPAKWANKRLCALNKSRRDPRLIYDGRDMPPOP 114
QY 119 HKVYALSRELCHWREGTGTQTEAEAPAAVPPQGGPPG-----PFLAHTH 163
DB 115 YKIY-----EVC-SNGF-----APTDSQPPEDYSFGAGEEEEEELQRLMPLSLSLTD 161
QY 164 AGLQAPGPLP-----APAGDKGDLLOAVQOQCLADHLLTASWGADPVPVKAPGE 213
DB 162 AVQSGPHMTFYSLLKEDVKWPTLQPTTLRPTLQPTLPVVLGPPADPSPLAPP-- 219
QY 214 GQGLPLTGACAGGPGLPAG--ELYGWAVETTPSGPOPAAATTGEEAAAPESHQAEVYL 271
DB 220 -----PGNPAGFRELLSEVLE-----PGPLPASL-----PPAGEQLPDL 254
QY 272 SPSPSACTAVQSPSGALDVTIMYKGRVTLQKVGHGHP--SCTFLYGGPDPVATDQ--- 327
DB 255 LISPHMLPTD-----LEIKFYGRGPRPRALTISNPHGCRFLFYS-----OLEATQOVEL 304
```


RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;	
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
RA	Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M.J., Udwin T.B., Toshitsuki S., Carninci P., Prange C.,	
RA	Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,	
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,	
RA	Krzywinski M.I., Skalska U., Smalios D.E., Schnierch A., Schein J.E.,	
RT	Jones S.J., Marra M.A.;	
RT	"Generation and initial analysis of more than 15,000 full-length human	
RL	and mouse cDNA sequences.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	
RL	[2]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RT	Dev. Dyn. 225:384-391(2002).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;	
RX		

Db 207 TPAYSTAPVHGAHAALP-----AEVATNOGSPLOQHITDHPFQDRTL 251
QY 287 GALDVITMYKGRVLOKVGHSCTFLYGPDPDAVRATDP---QQVAPSPAEPLDQK 342
Db 252 TEFEVTVYKGIKVSITLVNKP-----HGRITSRKHVSPGSLVDVLPSPMTIADQAV 306
QY 343 LRYTEELLRHVAPGLHLELRGPOLWARRMKCKVYVEVGPPGSGASPSPTACILPRNCDT 402
Db 307 VGEIHKLLRNLEGLVEVGGSGICRKQKCKRAFWMTMPISQPNQ-----IDKNDYC 362
QY 403 PIPDFRVFOELVEFRARORRGSPRTYILYFGQDLSAGRPKSKSLVVKLEPWLRCVHL 462
Db 363 ILHTLQOFAELTEFIERTKSPQYHIMWCLGELPDRVPMNKKFMQVIVPSMKLLH 422
QY 463 EGTQREGVSLDSDLDLCLSSANSYLDIECEFLMELEQ 501
Db 423 DMSYSTGASSLHSSSEINLEISDSLSSTNDVMVRLRELHE 461

RESULT 15

Q6PGZ7 PRELIMINARY; PRT; 492 AA.
AC Q6PGZ7
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein zgc:63500.
GN ORFNames=zgc:63500;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Jordan B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Zeeberg B., Moore T., Max S.I., Wang J., Hsieh L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerk A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC056772; AAH56772.1; -
DR ZFIN; ZDB-GENE-040426-1137; zgc:63500.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTRNRREGFCT.

DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Hypothetical protein.
SQ SEQUENCE 492 AA, 54876 MW, 98BF8FD3FELF589B CRC64;
Query Match 16.3%; Score 445.5; DB 2; Length 492;
Best Local Similarity 27.4%; Pred. No. 5.3e-17;
Matches 147; Conservative 74; Mismatches 187; Indels 129; Gaps 25;
QY 11 RVLFGWELGETSSGCGYEGLOWLDEARTCFRVPWKHFAR-KULSEADARLEKAWAVARGR 69
Db 7 RVLKFWLVSQVDNATFFGLVWLDRDAKQFIPWKGHATHTPQOESEENTLEKAWAVETGK 66
QY 70 WPPSSRGCGPPPEAEATAERAGKTNFRCAIARRRFRVLMRLDNGSD-PADPHKYVYALSREL 128
Db 67 Y---QEGVDEPDPAK-----WKAQLRCALNKSREENLYIDGTKEVPMPLKIY---DV 113
QY 129 CWREGFGTDTQTEAEAPAAVP-----PPQGGPPGPFLLAHTHAGLQAPGLPAPAG 177
Db 114 C-----DIPQTPSN-PGSVPTVETDGEDDIDPTPEPPPPYASH---GMSA-----SP-- 157
QY 178 DKGDLLOAVQOSCLADHLLTASWGADPVTKAPGEGQGLPLTGACAGGPGLPAGELYG 237
Db 158 -----LAAVW-----TPGGSVSLQP---SSCAPPPALP----- 183
QY 238 WAVETTPSPG-POPAALTTGEAAAP-----ESPHQAEPLYLSPSPSACTAVQEPSGAL 289
Db 184 -----PPALPPPAATVLPAAAAAPPANVWPKEPEDVE--MQPPP-----MEIQFTAL 230
QY 290 D-----VTIMYKGRVTLQKV-VGHP-SCITFLYGPDPDAVRATD--- 325
Db 231 DNLFITPETWISSLPMTDLEVPYRGKEVCPLTVSNPQGCRLFYGLDGLPIVNBELFG 290
QY 326 ---PQOVAFPSPAELPDQKQLRYTELLRHVAPGLHLELRGPOLWARRMKCKVYVEVG 382
Db 291 PVSLEQVRFPPTTEHIANDKQVFTSLLDVMDRGILLEVSGHDIYAVRLCQCKYVW---S 347
QY 383 PGASASPSPTACILPRNCDTPIDFRVFFQELVEFRARORRGSPRTYILYFGQDLSAGR 442
Db 348 GPCAPNPAPN-LIERQOKVKILFCLSFSLGVIHQRGSSDAPLDIHLFCFGEWPDGR 406
QY 443 PKEKSLVLVKLEPWLRCVHLEGTQREGVSSLDSDLDLCLSSANSYLDIECEFLMEL 499
Db 407 PREKLIWQVVPVVARMISEMPSGDCTRSFDSGVRQLQISIPD-IKDNIVTHLQKL 462

Search completed: May 6, 2005, 16:53:59

Job time : 183 secs